

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:48:41 ; Search time 2370 Seconds
(without alignments)

Sequence: 1 ggcattgacaggctaaatgc.....atctgctgattaggaagtat 300
8094.610 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_hg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298.4	99.5	121896	5	AL138713 Human DNA
2	143	47.7	281593	12	AC172613 Bos taurus
3	51.6	17.2	175961	6	AC124705 Mus muscu
4	51.6	17.2	184477	6	AC117638 Mus muscu
5	44.2	14.7	246893	12	AC129383 Rattus no
6	44.2	14.7	314404	12	AC130228 Rattus no
7	43.8	14.6	218961	12	AC153067 Bos tauru
8	43.6	14.5	4601	13	DMU11584 U11584 Drosophila
9	43.6	14.5	19517	13	DMU37541 U37541 Drosophila
10	43	14.3	173903	6	AC163669 Mus muscu
11	43	14.3	196140	12	AC084107 Mus muscu
12	42	14.0	154779	6	AC137708 Mus muscu
13	42	14.0	212071	6	AC136754 Mus muscu
14	42	14.0	223110	6	AC123752 Mus muscu
15	42	14.0	224009	12	AC131058 Mus muscu
16	41.8	13.9	168914	12	AC158640 Mus muscu
17	41.8	13.9	231134	6	AC111046 Mus muscu
18	41.6	13.9	237432	12	AC123182 Rattus no

c 19	41.6	13.9	242992	12	CR792426	CR792426
c 20	41.4	13.8	223752	12	AC094947	AC094947 Rattus no
c 21	41.4	13.8	253693	12	AC131839	AC131839 Rattus no
c 22	41.4	13.8	271927	12	AC098410	AC098410 Rattus no
c 23	40.8	13.6	23709	4	AP004980	AP004980 Lotus cor
c 24	40.8	13.6	158755	11	EX255970	EX255970 Zebrafish
c 25	40.8	13.6	165917	5	AL356073	AL356073 Human DNA
c 26	40.4	13.5	166494	12	AC135483	AC135483 Rattus no
c 27	40.4	13.5	183223	6	AC142454	AC142454 Mus muscu
c 28	40.4	13.5	186836	6	AC139360	AC139360 Mus muscu
c 29	40.4	13.5	192918	6	AC126460	AC126460 Mus muscu
c 30	40.4	13.5	213458	12	AC123390	AC123390 Rattus no
c 31	40.4	13.5	223392	12	AC095434	AC095434 Rattus no
c 32	40	13.3	110000	15	AE015927	Continuation (6 of
c 33	40	13.3	157168	5	AL137018	AL137018 Human DNA
c 34	40	13.3	160933	12	AC172335	AC172335 Bos tauru
c 35	40	13.3	181272	11	EX548256	EX548256 Zebrafish
c 36	40	13.3	192411	12	AC169312	AC169312 Bos tauru
c 37	39.8	13.3	4250	13	AB193599	AB193599 Plasmodiu
c 38	39.8	13.3	157011	12	AC140115	AC140115 Rattus no
c 39	39.8	13.3	161610	11	AL929174	AL929174 Zebrafish
c 40	39.8	13.3	185657	12	AC125630	AC125630 Rattus no
c 41	39.8	13.3	207949	11	EX005210	EX005210 Zebrafish
c 42	39.8	13.3	208337	12	AC126085	AC126085 Rattus no
c 43	39.6	13.2	46494	12	AC157740	AC157740 Xenopus t
c 44	39.6	13.2	110000	12	AC133155	Continuation (3 of
c 45	39.6	13.2	123109	11	EX511308	EX511308 Zebrafish
c 46	39.6	13.2	154673	6	AC166491	AC166491 Mus muscu
c 47	39.6	13.2	165154	6	AC121307	AC121307 Mus muscu
c 48	39.6	13.2	193188	6	AL844155	AL844155 Mouse DNA
c 49	39.6	13.2	200393	6	AC084288	AC084288 Mus muscu
c 50	39.6	13.2	229628	12	EX908768	EX908768 Mus muscu
c 51	39.4	13.1	48000	5	AY436325	AY436325 Homo sapi
c 52	39.4	13.1	60461	5	AF245699	AF245699 Homo sapi
c 53	39.4	13.1	110573	4	AC147741	AC147741 Medicago
c 54	39.4	13.1	159059	12	AC093000	AC093000 Homo sapi
c 55	39.4	13.1	160636	12	AC024991	AC024991 Homo sapi
c 56	39.4	13.1	163311	6	AC160632	AC160632 Mus muscu
c 57	39.4	13.1	204794	5	AC024897	AC024897 Homo sapi
c 58	39.4	13.1	223565	12	AC155681	AC155681 Bos tauru
c 59	39.4	13.1	225963	12	AC025153	AC025153 Homo sapi
c 60	39.4	13.1	270437	12	CR932018	CR932018 Danio rer
c 61	39.2	13.1	161092	11	CR318664	CR318664 Zebrafish
c 62	39.2	13.1	189185	12	AC173587	AC173587 Bos tauru
c 63	39.2	13.1	202967	12	AC175374	AC175374 Colobus g
c 64	39	13.0	661	7	BV413550	BV413550 S229P6408
c 65	39	13.0	15500	13	DQ241798	DQ241798 Scieropha
c 66	39	13.0	40852	13	CEFL1C1	Z54270 Caenorhabdi
c 67	39	13.0	125051	5	AL450164	AL450164 Human DNA
c 68	39	13.0	151319	5	AC108036	AC108036 Homo sapi
c 69	39	13.0	219318	12	AC140987	AC140987 Rattus no
c 70	38.8	12.9	110000	12	AC098516	Continuation (2 of
c 71	38.8	12.9	110000	12	AC114829	Continuation (3 of
c 72	38.8	12.9	175249	6	AC129603	AC129603 Mus muscu
c 73	38.8	12.9	201282	6	AC132450	AC132450 Mus muscu
c 74	38.8	12.9	289587	12	AC118337	AC118337 Rattus no
c 75	38.6	12.9	2224	2	C0489965	C0489965 Sequence
c 76	38.6	12.9	2224	2	C0490293	C0490293 Sequence
c 77	38.6	12.9	2224	2	C0490299	C0490299 Sequence
c 78	38.6	12.9	2224	2	C0491306	C0491306 Sequence
c 79	38.6	12.9	2224	2	C0492635	C0492635 Sequence
c 80	38.6	12.9	2224	2	C0492712	C0492712 Sequence
c 81	38.6	12.9	2224	2	C0492723	C0492723 Sequence
c 82	38.6	12.9	2224	2	C0493179	C0493179 Sequence
c 83	38.6	12.9	2224	2	C0495812	C0495812 Sequence
c 84	38.6	12.9	2224	2	C0496154	C0496154 Sequence
c 85	38.6	12.9	2224	2	C0496160	C0496160 Sequence
c 86	38.6	12.9	2224	2	C0497174	C0497174 Sequence
c 87	38.6	12.9	21691	13	CEY17D7C	AL021471 Caenorhab
c 88	38.6	12.9	137025	12	AC162158	AC162158 Rhinolph
c 89	38.6	12.9	147879	12	AC157887	AC157887 Rhinolph
c 90	38.6	12.9	211383	5	AC161278	AC161278 Pan trogl
c 91	38.6	12.9	227322	11	EX927261	EX927261 Zebrafish

```

c 92 38.4 12.8 110000 4 AP008209 045
c 93 38.4 12.8 110000 12 TANNIL_22
c 94 38.4 12.8 123340 4 AC151523
c 95 38.4 12.8 123340 4 AC151523
c 96 38.4 12.8 148894 4 AC121489
c 97 38.4 12.8 155293 12 AC154085
c 98 38.4 12.8 155633 12 AC025510
c 99 38.4 12.8 151804 5 AC020741
c 100 38.2 12.7 538 13 AY042101
c 101 38.2 12.7 45232 13 AF100657
c 102 38.2 12.7 99534 12 AC178452
c 103 38.2 12.7 127514 12 AC161985
c 104 38.2 12.7 155243 12 AC170794
c 105 38 12.7 19209 15 AF274444
c 106 38 12.7 171582 5 AC021558
c 107 38 12.7 183115 12 CT573496
c 108 38 12.7 184052 11 BX927369
c 109 38 12.7 198313 5 AC142314
c 110 38 12.7 214284 12 AC112544
c 111 37.8 12.6 12064 5 AF494057
c 112 37.8 12.6 41197 13 AF038614
c 113 37.8 12.6 67335 5 HS43683
c 114 37.8 12.6 156598 5 AC146185
c 115 37.8 12.6 168551 5 AC080007
c 116 37.8 12.6 168914 12 AC158640
c 117 37.8 12.6 173339 12 AC167778
c 118 37.8 12.6 194160 12 CR762444
c 119 37.8 12.6 194943 5 AC105252
c 120 37.8 12.6 231134 6 AC111046
c 121 37.8 12.6 245316 12 AC105887
c 122 37.6 12.5 30616 12 AC018027
c 123 37.6 12.5 101563 4 AC122543
c 124 37.6 12.5 138943 12 CR847865
c 125 37.6 12.5 163765 11 BX284684
c 126 37.6 12.5 165714 5 AC016919
c 127 37.6 12.5 171226 13 AC008144
c 128 37.6 12.5 175781 13 AC080891
c 129 37.6 12.5 176329 12 AC117968
c 130 37.6 12.5 177596 12 AC091945
c 131 37.6 12.5 178251 13 AE003737
c 132 37.6 12.5 213097 12 AC170271
c 133 37.6 12.5 225790 12 AC099213
c 134 37.6 12.5 230573 6 AC122883
c 135 37.6 12.5 244429 12 AC155061
c 136 37.4 12.5 34368 13 CEW06D12
c 137 37.4 12.5 35913 4 AC149414
c 138 37.4 12.5 63233 5 AL359504
c 139 37.4 12.5 157929 5 AL593844
c 140 37.4 12.5 160082 5 AP003695
c 141 37.4 12.5 164485 12 AC025340
c 142 37.4 12.5 164768 12 AC140766
c 143 37.4 12.5 169660 5 AP003547
c 144 37.4 12.5 189473 12 AC022950
c 145 37.4 12.5 195400 12 AC119022
c 146 37.4 12.5 200458 12 AC019237
c 147 37.4 12.5 201524 12 AC178557
c 148 37.4 12.5 203860 12 AC162535
c 149 37.4 12.5 208961 5 AC090819
c 150 37.4 12.5 210649 12 AC172121

```

ALIGNMENTS

```

RESULT 1
AL138713 121896 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP11-552M6 on chromosome
DEFINITION 13q21.31-22.2 Contains part of the KLF12 gene for Kruppel-like
factor 12, complete sequence.
ACCESSION AL138713
VERSION AL138713.12 GI:14625506
KEYWORDS HTG; KLF12.

```

Continuation (46 o
Continuation (23 o
AC151523 Medicago
AC121489 Oryza sat
AC154085 Rhinoph
AC025510 Homo sapi
AC020741 Homo sapi
AY042101 Neoelebo
AF100657 Caenorhab
AC178452 Strongylo
AC161985 Bos tauru
AC170794 Bos tauru
AF274444 Carsonell
AC021558 Homo sapi
CT573496 Danio rer
BX927369 Zebrafish
AC142314 Pan trogl
AC112544 Rattus no
AC095474 Rattus no
AF494057 Homo sapi
AF038614 Caenorhab
AL49212 Homo sapi
AC146185 Pan trogl
AC080007 Homo sapi
AC158640 Mus muscu
AC167778 Bos tauru
CR762444 Danio rer
AC105252 Homo sapi
AC111046 Mus muscu
AC105887 Rattus no
AC018027 Drosophil
AC122543 Genomic s
CR847865 Danio rer
BX284684 Zebrafish
AC016919 Homo sapi
AC008144 Drosophil
AC080891 Drosophil
AC117968 Rattus no
AC091945 Homo sapi
AE003737 Drosophil
AC170271 Bos tauru
AC099213 Rattus no
AC122883 Mus muscu
AC155061 Bos tauru
282073 Caenorhabdi
AC149414 Phakopsor
AL359504 Human DNA
AL593844 Human DNA
AP003695 Homo sapi
AC025340 Homo sapi
AC140766 Rattus no
AP003547 Homo sapi
AC022950 Homo sapi
AC119022 Rattus no
AC019237 Homo sapi
AC178557 Strongylo
AC162535 Bos tauru
AC090819 Homo sapi
AC172121 Bos tauru

Source

FEATURES

Location/Qualifiers

1. 121896
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="q21.31-22.2"
/clone="RP11-552M6"
/clone_lib="RP11-11.2"
100
/note="Clone right end: RP11-474L7"
join(complement(AL139036.12:11189.11255),
complement(AL159972.8:109592.109681),
complement(AL159972.8:11438.11984),
complement(104572.104707),complement(56358.56420),
complement(6788.6945),
complement(AL160032.14:158733.168315))
/gene="KLF12"
/locus_tag="RP11-474L7.5-001"
join(complement(AL139036.12:11189.11255),
complement(AL159972.8:109592.109681),
complement(AL159972.8:11438.11984),
complement(104572.104707),complement(56358.56420),
complement(6788.6945),
complement(AL160032.14:158733.168315))
/gene="KLF12"
/product="Kruppel-like factor 12"
/note="match: ESTs: AA007205 AA349967 AA513497 AA524929
AI373658 AI628470 AL701841 AV728478 BG834258 BG910766
F12639 R23424 R23427 R23428 T74529
match: cDNAs: AF113122 AF161471 AF330041 AJ243274 BC019680
Y14295"
join(complement(AL139036.12:11189.11221),
join(complement(AL159972.8:109592.109681),

misc_feature

gene

mRNA

CDS

45968 AAACAGCAGCTCCCAATTTGTCCTAATACTAGCGAAATCCATCATGAAGAAATG--TAAA 46025
 124 TGCTTAATTTTCAATAAGAGTACACAGAGTGTGCAACATGCTGTTAAATAACCCGAC 183
 46026 AGCTATTTTAAATAAAGTACACAGAGTATTGCAACATCTCTGTTAAACAGCAAAAC 46085
 184 AAATCTCAATCACTATAGCTGTAGTAGAGTCAATTCCTGCAAGATCCAGAGTAACCACT 243
 46086 AAATTCACCTCATACAGCTGTAGTAGAGTCTGCTGTTAGAAATTCAGAGTAACTGAT 46145
 244 ATTTTGGAAATGCAATGTTGAACCGACCATCACTAATTAATCTGCTGATTAGGAAGTAT 300
 46146 ATTTTGGAAACTGGAATGATTAACCTCCTAGACAGTATCTGCTGATTAGGAAGCAT 46202

RESULT 3

AG124705/c
 LOCUS AC124705 175961 bp DNA linear ROD 27-NOV-2003
 DEFINITION Mus musculus BAC clone RP24-68G21 from chromosome 14, complete sequence.

AC124705

AC124705.4 GI:37951470

HTG

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 175961)

Levy, A., Haakenson, W., Schatzkammer, K. and Bielicki, L.

The sequence of Mus musculus BAC clone RP24-68G21

2 (bases 1 to 175961)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 175961)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 175961)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (01-NOV-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 175961)

Wilson, R.K.

Direct Submission

Submitted (24-OCT-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 175961)

Wilson, R.

Direct Submission

Submitted (27-NOV-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Oct 24, 2003 this sequence version replaced gi:24462442.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0068G21

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC117638.

FEATURES

source

Location/Qualifiers

1..175961
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="14"
 /map="14"
 /clone="RP24-68G21"
 /clone_lib="RPCI-24"
 /rpt_family="ERVK"
 699..779
 /rpt_family="B4"
 799..951
 /rpt_family="B4"
 1935..2049
 /rpt_family="MIR"
 3410..3612
 /rpt_family="B2"
 4481..4664
 /rpt_family="L1"
 4783..5177
 /rpt_family="ERV1"
 6224..6294
 /rpt_family="L1"
 8387..8441
 /rpt_family="ERV1"
 11635..12030
 /rpt_family="L1"
 12282..12493
 /rpt_family="B2"
 12523..12583
 /rpt_family="MER1_type"
 12645..12780
 /rpt_family="L1"
 13231..13338
 /rpt_family="Alu"
 14632..14820
 /rpt_family="B4"
 14881..14998
 /rpt_family="Alu"
 15132..15790
 /rpt_family="L1"
 15856..16126
 /rpt_family="MaLR"
 18218..18615
 /rpt_family="MaLR"
 18865..19213
 /rpt_family="MaLR"
 19661..20020
 /rpt_family="L1"
 20040..20129
 /rpt_family="L1"
 21275..21402

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region	/rpt_family="Alu"	21997..22141
repeat_region	/rpt_family="MaLR"	23550..23946
repeat_region	/rpt_family="MaLR"	26897..26863
repeat_region	/rpt_family="L1"	27825..27984
repeat_region	/rpt_family="B2"	29411..29607
repeat_region	/rpt_family="ERVK"	31300..31876
repeat_region	/rpt_family="L1"	31943..32107
repeat_region	/rpt_family="B4"	33292..33385
repeat_region	/rpt_family="L1"	35024..35221
repeat_region	/rpt_family="B2"	35207..35467
repeat_region	/rpt_family="L1"	35468..35589
repeat_region	/rpt_family="L1"	36177..36579
misc_feature	/rpt_family="MaLR"	45951..46588
repeat_region	/note="Sequence de	47048..47115
repeat_region	/rpt_family="L1"	47115..47807
repeat_region	/rpt_family="L1"	49387..49573
repeat_region	/rpt_family="B2"	51674..52329
repeat_region	/rpt_family="L1"	52330..53173
repeat_region	/rpt_family="L1"	54498..54652
repeat_region	/rpt_family="B4"	54840..54992
repeat_region	/rpt_family="L1"	55487..55791
repeat_region	/rpt_family="B4"	57343..57541
repeat_region	/rpt_family="B2"	57678..57951
repeat_region	/rpt_family="B4"	58729..58950
repeat_region	/rpt_family="MaLR"	62188..62448
repeat_region	/rpt_family="B4"	63148..63422
repeat_region	/rpt_family="MaLR"	65214..65481
repeat_region	/rpt_family="L1"	65767..65935
repeat_region	/rpt_family="B4"	66245..66353
repeat_region	/rpt_family="Alu"	68642..68729
repeat_region	/rpt_family="B4"	68920..68997
repeat_region	/rpt_family="B4"	69011..69155
repeat_region	/rpt_family="Alu"	69382..69672
repeat_region	/rpt_family="L1"	70522..71262
repeat_region	/rpt_family="L1"	71271..71381
repeat_region	/rpt_family="L1"	71381..71381

Query Match

17.2%; Score 51.6; DB 6; Length 175961;

Best Local Matches	
Qy	14
Db	14
Qy	
Db	14
Qy	
Db	14
RESULT 4	
AC117638//	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

RESULT 4	REFERENCE
AC117638/4	AUTHORS
LOCUS	TITLE
DEFINITION	JOURNAL
ACCESSION	AUTHORS
VERSION	
KEYWORDS	
SOURCE	
ORGANISMS	

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS	REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS
1. H. J. ... 2. ... 3. ... 4. ... 5. ... 6. ... 7. ... 8. ... 9. ... 10. ... 11. ... 12. ... 13. ... 14. ... 15. ... 16. ... 17. ... 18. ... 19. ... 20. ... 21. ... 22. ... 23. ... 24. ... 25. ... 26. ... 27. ... 28. ... 29. ... 30. ... 31. ... 32. ... 33. ... 34. ... 35. ... 36. ... 37. ... 38. ... 39. ... 40. ... 41. ... 42. ... 43. ... 44. ... 45. ... 46. ... 47. ... 48. ... 49. ... 50. ... 51. ... 52. ... 53. ... 54. ... 55. ... 56. ... 57. ... 58. ... 59. ... 60. ... 61. ... 62. ... 63. ... 64. ... 65. ... 66. ... 67. ... 68. ... 69. ... 70. ... 71. ... 72. ... 73. ... 74. ... 75. ... 76. ... 77. ... 78. ... 79. ... 80. ... 81. ... 82. ... 83. ... 84. ... 85. ... 86. ... 87. ... 88. ... 89. ... 90. ... 91. ... 92. ... 93. ... 94. ... 95. ... 96. ... 97. ... 98. ... 99. ... 100. ...	1. H. J. ... 2. ... 3. ... 4. ... 5. ... 6. ... 7. ... 8. ... 9. ... 10. ... 11. ... 12. ... 13. ... 14. ... 15. ... 16. ... 17. ... 18. ... 19. ... 20. ... 21. ... 22. ... 23. ... 24. ... 25. ... 26. ... 27. ... 28. ... 29. ... 30. ... 31. ... 32. ... 33. ... 34. ... 35. ... 36. ... 37. ... 38. ... 39. ... 40. ... 41. ... 42. ... 43. ... 44. ... 45. ... 46. ... 47. ... 48. ... 49. ... 50. ... 51. ... 52. ... 53. ... 54. ... 55. ... 56. ... 57. ... 58. ... 59. ... 60. ... 61. ... 62. ... 63. ... 64. ... 65. ... 66. ... 67. ... 68. ... 69. ... 70. ... 71. ... 72. ... 73. ... 74. ... 75. ... 76. ... 77. ... 78. ... 79. ... 80. ... 81. ... 82. ... 83. ... 84. ... 85. ... 86. ... 87. ... 88. ... 89. ... 90. ... 91. ... 92. ... 93. ... 94. ... 95. ... 96. ... 97. ... 98. ... 99. ... 100. ...

TITLE	JOURNAL	REFERENCE	AUTHORS
1. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
2. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
3. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
4. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
5. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
6. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
7. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
8. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
9. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.
10. The effect of the concentration of the solution on the rate of the reaction	Journal of Physical Chemistry	1950	Smith, J. D.


```

Query Match      17.2%; Score 51.6; DB 6; Length 184477;
Best Local Similarity 64.4%; Pred. No. 7.7e-06;
Matches 112; Conservative 0; Mismatches 54; Indels 8; Gaps 2;

QY 64 AAACAGCTAACTCCCAATGTTACCTAACTAGGAGGAGTTTATCATGAAGAAATGTTTAAA 123
Db 110061 ACATAGCCAGCTCCCAATGTTATCATACAGGAATAATCT-----ATTAGAAATGTTTGT 110006

QY 124 TGCTTAATTTTCAAAATGAAGTAACACGAGTGTGTGCAACATGCTGTTAAATACCCGAC 183
Db 110005 TGCTTTATTTT-----ACAGAAACTTAACCAAAATGTTGCAACATTCGGCTAAGTAGTAGAC 109950

QY 184 AAACCTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCACAGTA 237
Db 109949 AAACCTAAATCTCAGCAGCAATCCCTCTTACGATTCGTCAGGAATGCTAGAGTA 109896

RESULT 5
AC129383
LOCUS      Rattus norvegicus clone CH230-36A10, *** SEQUENCING IN PROGRESS
DEFINITION      ***, 3 unordered pieces.
ACCESSION      AC129383
VERSION        AC129383.5 GI:24818055
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muridae; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS        Muzny, D., Maric, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
               Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
               Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
               Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
               Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
               Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
               Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, A.,
               Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, X., Chen, Z., Chu, J.,
               Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
               Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
               Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
               Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
               Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
               Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
               Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
               Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
               Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
               Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
               Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
               Hollins, B., Howells, S., Hulyk, S., Hume, J., Idliebird, D., Jackson, A.,
               Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
               Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
               Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
               Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
               Lorensuhwa, L., Louissegh, H., Lozada, R.J., Lu, X., Ma, J.,
               Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
               Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
               Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
               Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
               Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
               Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
               Nwakeleneh, O., Okwunonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
               Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
               Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
               Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
               Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
               Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
               Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
               Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,
               Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,
               Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
               Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

```

```

Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 246893)
Worley, K.C.
Direct Submission
Submitted (29-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246893)
Rat Genome Sequencing Consortium.
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23096176.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GEEP
Center clone name: CH230-36A10
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 225900 bases at least Q40
Consensus quality: 229196 bases at least Q30
Consensus quality: 231050 bases at least Q20
Estimated insert size: 230916; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 243625: contig of 243625 bp in length
* 243626 243725: gap of unknown length
* 243726 244798: contig of 1073 bp in length
* 244799 244899: gap of unknown length
* 244899 246893: contig of 1995 bp in length.
* Location/Qualifiers
* 1..246893
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clones="CH230-36A10"
* 1..1304
* /note="wgs_contig"
FEATURES
source
misc_feature

```

```

misc_feature 3950..5173
      /note="wgs_contig"
misc_feature 140573..142683
      /note="wgs_contig"
misc_feature complement(233360..234110)
      /note="clone_boundary"
      clone_end:17
      site:ECORI
      end_sequence:BH299246"
misc_feature 234187..235820
      /note="wgs_end_extension"
      clone_end:17"
misc_feature 242264..243625
      /note="wgs_end_extension"
      clone_end:17"
gap 243626..243725
      /estimated_length=unknown
gap 244799..244898
      /estimated_length=unknown

ORIGIN
Query Match 14.7%; Score 44.2; DB 12; Length 246893;
Best Local Similarity 61.8%; Pred. No. 0.0022;
Matches 107; Conservative 0; Mismatches 58; Indels 8; Gaps 2;

Qy 64 AAACGCTAATCCCAATGTGACCTAATAGGAGGAGTTATCATGAAGAATGTTTAAA 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135556 ACACAGCCCATTTCCCAATGTATATACACAGGAAA-----TTCATGAGGAATGTTTCA 135611

Qy 124 TGCCTTAATTTCAATATAGAGTACACAGAGTGTTCACATGCTGTTAATATACCCGAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135612 TGCCTTTGTTTAGAA-----AAACTAACCAAGTGTGCAACACTTTGTTTAAGTCGTAGAC 135667

Qy 184 AAATCTCAATCACTATAGTGTAGTAGAGTGCATTCGCAAGGATCCAGAGT 236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135668 AAACGGAATCTCTAGCAATCTCTTTTACCATTCTGCAGAAATCTGGACT 135720

RESULT 6
AC130228/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-53K23, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
ACCESSION AC130228
VERSION AC130228.4 GI:24635276
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 314404)
Muzny D.Marie, Metzker M.Lee, Abramzon S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalbechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogue M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,

```

```

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhs, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McJeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfaunkoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,
Sneid, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (Bases 1 to 314404)
Direct Submission
Submitted (09-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (Bases 1 to 314404)
Rat Genome Sequencing Consortium.
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23264801.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCIZ
Center clone name: CH230-53K23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 242417 bases at least Q40
Consensus quality: 244920 bases at least Q30
Consensus quality: 246473 bases at least Q20
Estimated insert size: 246166; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

```

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 312945: contig of 312945 bp in length
 * 312946 313045: gap of unknown length
 * 313046 314404: contig of 1359 bp in length.

FEATURES

```

source
    1..314404
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-53K23"
    1..1916
        /note="wgs end extension"
        /clone_end:Sp6"
        2588..4415
        /note="wgs end extension"
        /clone_end:Sp6"
        complement(5720..6427)
        /note="clone boundary"
        /clone_end:Sp6
        site:
            end sequence:BH323709"
            242173..243018
            /note="clone boundary"
            /clone_end:T7
        site:
            end sequence:BH323707"
            243109..244516
            /note="wgs end extension"
            /clone_end:T7"
            245081..246984
            /note="wgs end extension"
            /clone_end:T7"
            310363..312945
            /note="wgs end extension"
            /clone_end:T7"
            312946..313045
            /estimated_length=unknown

```

ORIGIN

```

Query Match      14.7%; Score 44.2; DB 12; Length 314404;
Best Local Similarity 61.8%; Pred. No. 0.0022;
Matches 107; Conservative 0; Mismatches 56; Indels 8; Gaps 2;

Qy 64 AACAGCTAACTCCCAATTTACTACTAGGGGAGTTTATCATGAGAAATGTTTAA 123
Db 41164 ACACAGCCATCCCAATGTTATATACAGGAAA---TTCATGAGGAATGTTTCA 41109

Qy 124 TGCCTTAATTTCAATAGAGTACACAGATGTTGCAACATGCTGTTAATATACCCGAC 183
Db 41108 TGCCTTTGTTTAGAA---AAACTAACCAAGTGTGCAACACTTTGTTAAGTCGTAGAC 41053

Qy 184 AAATCTCAATCATATAGCTAGTAGTGCATTCGACAGGATCCAGAGT 236
Db 41052 AAACGGAATCTCTGTAGCAATTCCTTTTACCATTCGACAAATATCTGGACT 41000

```

RESULT 7

```

AC153067/c
LOCUS AC153067 218961 bp DNA linear HTG 01-JUL-2005
DEFINITION Bos taurus clone CH240-1302, *** SEQUENCING IN PROGRESS ***, 31
unordered pieces.
ACCESSION AC153067
VERSION AC153067.3 GI:68226955
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS

1 (bases 1 to 218961)
 Muzny,D,Marie,, Metzker,M, Lee,, Abramzon,S,, Adams,C,, Alder,J,, Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,, Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,, Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,, Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,, Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,, Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,, Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,, Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,, Davila,M,L,, Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,, Delgado,O,, Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,, Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,, Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falls,T,, Fan,G,, Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,, Fraser,C,M,, Gabisi,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,, Gebregiorgis,E,, Geer,K,, Gill,R,, Grady,M,, Guerra,W,, Guevara,W,, Gunaratne,P,, Haaland,W,, Hamil,C,, Hamilton,C,, Hamilton,K,, Harvey,Y,, Havlak,P,, Hawes,A,, Henderson,N,, Hernandez,J,, Hernandez,R,, Hines,S,, Hladun,S,L,, Hodgson,A,, Hogues,M,, Hollins,B,, Howells,S,, Hulyk,S,, Hume,J,, Idlebird,D,, Jackson,A,, Jackson,L,, Jacob,L,, Jiang,H,, Johnson,B,, Johnson,R,, Jolivet,A,, Karpachy,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Kovat,C,, Kowis,C,, Kraft,C,L,, Lebow,H,, Levan,J,, Lewis,L,, Li,Z,, Liu,J,, Liu,J,, Liu,W,, Liu,Y,, London,P,, Longacre,S,, Lopez,J,, Lorensuhewa,L,, Loulseghe,H,, Lozado,R,J,, Lu,X,, Ma,J,, Maheshwari,M,, Mahindartne,M,, Mahmoud,M,, Malloy,K,, Mangum,A,, Mangun,B,, Napua,P,, Martin,K,, Martin,R,, Martinez,E,, Mawhney,S,, McLeod,M,P,, McNeill,T,Z,, Meenen,E,, Milosavljevic,A,, Miner,G,, Minja,E,, Montemayor,J,, Moore,S,, Morgan,M,, Morris,K,, Morris,S,, Munidasa,M,, Murphy,M,, Nair,L,, Nankervis,C,, Neal,D,, Newton,N,, Nguyen,N,, Norris,S,, Nwaokemele,O,, Okwuonu,G,, Olarnpunsagoon,A,, Pal,S,, Parks,K,, Pasternak,S,, Paul,H,, Perez,A,, Perez,L,, Pfannkuch,C,, Plopper,F,, Poindexter,A,, Popovic,D,, Primus,E,, Pu,L,L,, Puazo,M,, Quiroz,J,, Rachlin,E,, Reeves,K,, Regier,M,A,, Reigh,R,, Reilly,B,, Reilly,M,, Ren,Y,, Reuter,M,, Richards,S,, Riggs,F,, Rives,C,, Rodkey,T,, Rojas,A,, Rose,M,, Rose,R,, Ruiz,S,J,, Sanders,W,, Savery,G,, Scherer,S,, Scott,G,, Shatsman,S,, Shen,H,, Shetty,J,, Shvartsbeyn,A,, Sisson,I,, Sitter,C,D,, Smajs,D,, Sneed,A,, Sodergren,E,, Song,X,-Z,, Sorelle,R,, Sosa,J,, Steimle,M,, Strong,R,, Sutton,A,, Svatek,A,, Taber,P,, Taylor,C,, Taylor,T,, Thomas,N,, Thomas,S,, Tingey,A,, Trejos,Z,, Usmani,K,, Valas,R,, Vera,V,, Villaseana,D,, Waldron,L,, Walker,B,, Wang,J,, Wang,Q,, Wang,S,, Warren,J,, Warren,R,, Wei,X,, White,F,, Williams,G,, Willson,R,, Wleczyk,R,, Wooden,H,, Worley,K,, Wright,D,, Wright,R,, Wu,J,, Yakub,S,, Yen,J,, Yoon,L,, Yoon,V,, Yu,F,, Zhang,J,, Zhou,J,, Zhou,X,, Zhao,S,, Dunn,D,, von Niederhausern,A,, Weiss,R,, Smith,D,R,, Holt,R,A,, Smith,H,O,, Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jun 26, 2005 this sequence version replaced gi:58037848.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FAOQ
Center clone name: CH240-13J2
----- Summary Statistics

Consensus quality: 203243 bases at least Q40
Consensus quality: 206548 bases at least Q30
Consensus quality: 209635 bases at least Q20
Estimated insert size: 212382; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation table.

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 12238: contig of 12238 bp in length
* 12239 12515: gap of 277 bp
* 12516 14308: contig of 1793 bp in length
* 14309 14408: gap of unknown length
* 14409 23603: contig of 9195 bp in length
* 23604 23653: gap of 50 bp
* 23654 28728: contig of 5075 bp in length
* 28729 38011: contig of 9233 bp in length
* 38012 38061: gap of 50 bp
* 38062 41982: contig of 3921 bp in length
* 41983 42032: gap of 50 bp
* 42033 47531: contig of 5499 bp in length
* 47532 47581: gap of 50 bp
* 47582 82655: contig of 35074 bp in length
* 82656 82705: gap of 50 bp
* 82706 84911: contig of 2206 bp in length
* 84912 85011: gap of unknown length
* 85012 86758: contig of 1747 bp in length
* 86759 86808: gap of 50 bp
* 86809 103654: contig of 16846 bp in length
* 103655 103704: gap of 50 bp
* 103705 108056: contig of 4352 bp in length
* 108057 108156: gap of unknown length
* 108157 110213: contig of 2057 bp in length
* 110214 114233: gap of 1210 bp
* 11424 113681: contig of 2258 bp in length
* 113682 113781: gap of unknown length
* 113782 115505: contig of 1724 bp in length
* 115506 115555: gap of 50 bp
* 115556 122394: contig of 6839 bp in length
* 122395 122494: gap of unknown length
* 12495 124304: contig of 1810 bp in length
* 124305 124354: gap of 50 bp
* 124355 144012: contig of 19658 bp in length
* 144013 144924: gap of 912 bp
* 144925 150751: contig of 5827 bp in length
* 150752 150801: gap of 50 bp
* 150802 154839: contig of 4038 bp in length
* 154840 154889: gap of 50 bp

```

```

* 154890 161537: contig of 6648 bp in length
* 161538 161587: gap of 50 bp
* 161588 171569: contig of 9982 bp in length
* 171570 171619: gap of 50 bp
* 171620 181844: contig of 10225 bp in length
* 181845 181894: gap of 50 bp
* 181895 184328: contig of 2434 bp in length
* 184329 184901: gap of 573 bp
* 184902 199708: contig of 14807 bp in length
* 199709 199758: gap of 50 bp
* 199759 210817: contig of 11059 bp in length
* 210818 210917: gap of unknown length
* 210918 211927: contig of 1010 bp in length
* 211928 212027: gap of unknown length
* 212028 213436: contig of 1409 bp in length
* 213437 213536: gap of unknown length
* 213537 214731: contig of 1195 bp in length
* 214732 214831: gap of unknown length
* 214832 217364: contig of 2533 bp in length
* 217365 217464: gap of unknown length
* 217465 218961: contig of 1497 bp in length.

```

FEATURES

```

Source
1. 218961
   /organism="Bos taurus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9913"
   /clone="CH240-13J2"
   12239..12515
   /estimated_length=277
   14309..14408
   /estimated_length=unknown
   23604..23653
   /estimated_length=50
   28729..28778
   /estimated_length=50
   38012..38061
   /estimated_length=50
   41983..42032
   /estimated_length=50
   47532..47581
   /estimated_length=50
   82656..82705
   /estimated_length=50
   84912..85011
   /estimated_length=unknown
   86759..86808

```

```

Query Match      14.6%; Score 43.8; DB 12; Length 218961;
Best Local Similarity 51.8%; Pred. No. 0.0031;
Matches 99; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 5 TTGACAGGCTAAATGCTAAGTGACGTACGATGAAGACTTGAATATCTTCATTGGAATA 64
Db 40349 TTGAGCATATAAATGACAAAGCAATAATATGTTGAAATAATATTTATATTGTCATTT 40290

Qy 65 AACAGCTAACTCCCAATTTGACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAAT 124
Db 40289 ATAAAGGCTCTCTCGAATATTTCTCACTATTTATGTTTTATAGTGAAGTGAATTT 40230

Qy 125 GCTTAATTTTCAATAAAGAGTAACCAAGAGTGTGTGCAACATGCTGTGTTAAATACCCGACA 184
Db 40229 GCTTATTATAGTAAGATTATATAACATATTGTTGCAAAAGCAATTCATTTCRAACCAC 40170

Qy 185 AACTTCAATCA 195
Db 40169 AAGTTCAATCA 40159

RESULT 8
DMU11584
LOCUS
DEFINITION Drosophila melanogaster Oregon-R mitochondrial A+T region.
ACCESSION U11584

```



```

VERSION      U11584.1  GI:508826
KEYWORDS     mitochondrial DNA; A+T region; tandem repeats.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 4601)
AUTHORS      Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE        Sequence, organization, and evolution of the A+T region of
              Drosophila melanogaster mitochondrial DNA
JOURNAL      Mol. Biol. Evol. 11 (3), 523-538 (1994)
PUBMED      8015445
REFERENCE    2 (bases 1 to 4601)
AUTHORS      Kaguni,L.S.
TITLE        Direct Submission
JOURNAL      Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D. Dept. of
              Biochemistry, Michigan State University, East Lansing, MI,
              48824-1318, USA
FEATURES     source
              1..4601
               /organism="Drosophila melanogaster"
               /organella="mitochondrion"
               /mol_type="genomic DNA"
               /strain="Oregon-R"
               /db_xref="taxon:7227"
               /dev_stage="embryo"
               650..1022
               /note="repeat I-A"
               /rpt_type=tandem
               1023..1360
               /note="repeat I-B1"
               /rpt_type=tandem
               1361..1705
               /note="repeat I-C/A"
               /rpt_type=tandem
               1706..2043
               /note="repeat I-B2"
               /rpt_type=tandem
               2044..2388
               /note="repeat I-C"
               /rpt_type=tandem
               2491..2511
               /note="deoxythymidylate stretch"
               2512..2648
               /rpt_type=tandem
               2649..3112
               /note="repeat II-A"
               /rpt_type=tandem
               3113..3576
               /note="repeat II-B1"
               /rpt_type=tandem
               3577..4040
               /note="repeat II-B2"
               /rpt_type=tandem
               4041..4504
               /note="repeat II-C"
               /rpt_type=tandem
               complement(4565..4585)
               /note="deoxythymidylate stretch"
ORIGIN
Query Match      14.5%; Score 43.6; DB 13; Length 4601;
Best Local Similarity 58.5%; Pred. No. 0.0038;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy      29  TACGATGAAGACTTGAATATCTTCATTTTGAATAAACACGTAACCCCAATTGTACC 88
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1494 TATATATTAACCTTAAACAAAATATTTTATTAATAATAATAATAATAATAATA 1553
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      89  TAACAGGGAGTTTATCATCAAGAAATGTTTAATGCTTAATTTTCAATAAAGAGTAA 148
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1554 TATATTAACCTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1613
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy      149 CCAGAGTGT 158
Db      1614 GAAATATTTT 1623

RESULT 9
DMU37541
LOCUS       19517 bp      DNA      circular INV 02-MAR-2001
DEFINITION Drosophila melanogaster complete mitochondrial genome.
ACCESSION  U37541
VERSION     U37541.1  GI:1166529
KEYWORDS   .
SOURCE     mitochondrion Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 12511 to 12682)
AUTHORS    Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and
            Wolstenholme,D.R.
TITLE      Drosophila mitochondrial DNA: a novel gene order
JOURNAL    Nucleic Acids Res. 10 (21), 6619-6637 (1982)
PUBMED     6294611
REFERENCE  2 (bases 5269 to 5695)
AUTHORS    Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.
TITLE      Transfer RNA genes in Drosophila mitochondrial DNA: related 5'
            flanking sequences and comparisons to mammalian mitochondrial tRNA
            genes
JOURNAL    Nucleic Acids Res. 11 (8), 2411-2425 (1983)
PUBMED     6304852
REFERENCE  3 (bases 404 to 5272)
AUTHORS    de Bruijn,M.H.
TITLE      Drosophila melanogaster mitochondrial DNA, a novel organization and
            genetic code
JOURNAL    Nature 304 (5923), 234-241 (1983)
PUBMED     6408489
REFERENCE  4 (bases 804 to 1778)
AUTHORS    Satta,Y., Ishiwa,H. and Chigusa,S.I.
TITLE      Analysis of nucleotide substitutions of mitochondrial DNAs in
            Drosophila melanogaster and its sibling species
JOURNAL    Mol. Biol. Evol. 4 (6), 638-650 (1987)
PUBMED     2832697
REFERENCE  5 (bases 5268 to 13619)
AUTHORS    Garesse,R.
TITLE      Drosophila melanogaster mitochondrial DNA: gene organization and
            evolutionary considerations
JOURNAL    Genetics 118 (4), 649-663 (1988)
PUBMED     3130291
REFERENCE  6 (bases 441 to 2967)
AUTHORS    Satta,Y. and Takahata,N.
TITLE      Evolution of Drosophila mitochondrial DNA and the history of the
            melanogaster subgroup
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)
PUBMED     2124697
REFERENCE  7 (bases 14215 to 14512)
AUTHORS    Ballard,J.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and
            Atkinson,P.W.
TITLE      Evidence from 12S ribosomal RNA sequences that onychophorans are
            modified arthropods
JOURNAL    Science 258 (5086), 1345-1348 (1992)
PUBMED     1455227
REFERENCE  8 (bases 14917 to 19517)
AUTHORS    Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE      Sequence organization, and evolution of the A+T region of
            Drosophila melanogaster mitochondrial DNA
JOURNAL    Mol. Biol. Evol. 11 (3), 523-538 (1994)
PUBMED     8015445
REFERENCE  9 (bases 1 to 408; 13319 to 19517)
AUTHORS    Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE      Drosophila melanogaster mitochondrial DNA: completion of the
            nucleotide sequence and evolutionary comparisons
JOURNAL    Insect Mol. Biol. 4 (4), 263-278 (1995)
PUBMED

```



```

Contact: nisc_mouse@nhgri.nih.gov
----- Project Information
Center project name: qc
Center clone name: 343A21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 179429 bases at least Q40
Consensus quality: 185220 bases at least Q30
Consensus quality: 188059 bases at least Q20
Insert size: 196000; agarose-fp
Insert size: 192940; sum-of-contigs
Quality coverage: 4.03x in Q20 bases; agarose-fp
Quality coverage: 4.09x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1293: contig of 1293 bp in length
* 1294: gap of unknown length
* 1394: 2711: contig of 1318 bp in length
* 2712: 2811: gap of unknown length
* 2812: 3945: contig of 1134 bp in length
* 3946: 4045: gap of unknown length
* 4046: 5332: contig of 1287 bp in length
* 5333: 5432: gap of unknown length
* 5433: 7745: contig of 2313 bp in length
* 7746: 7845: gap of unknown length
* 7846: 10021: contig of 2176 bp in length
* 10022: 10121: gap of unknown length
* 10122: 13447: contig of 3326 bp in length
* 13448: 13547: gap of unknown length
* 13548: 16926: contig of 3379 bp in length
* 16927: 17026: gap of unknown length
* 17027: 21840: contig of 4814 bp in length
* 21841: 21940: gap of unknown length
* 21941: 26810: contig of 4870 bp in length
* 26811: 26910: gap of unknown length
* 26911: 31227: contig of 4317 bp in length
* 31228: 31327: gap of unknown length
* 31328: 36553: contig of 5226 bp in length
* 36554: 36654: gap of unknown length
* 36654: 42253: contig of 5600 bp in length
* 42254: 42354: gap of unknown length
* 42354: 46755: contig of 4402 bp in length
* 46756: 46855: gap of unknown length
* 46856: 51326: contig of 4471 bp in length
* 51327: 51427: gap of unknown length
* 51427: 55184: contig of 3758 bp in length
* 55185: 55284: gap of unknown length
* 55285: 60367: contig of 5083 bp in length
* 60368: 60467: gap of unknown length
* 60468: 65557: contig of 5090 bp in length
* 65558: 65658: gap of unknown length
* 65658: 70384: contig of 4727 bp in length
* 70385: 70484: gap of unknown length
* 70485: 75501: contig of 5017 bp in length
* 75502: 75602: gap of unknown length
* 75602: 81571: contig of 5969 bp in length
* 81571: 81670: gap of unknown length
* 81671: 86308: contig of 4638 bp in length
* 86309: 86408: gap of unknown length
* 86409: 92383: contig of 5975 bp in length
* 92384: 92483: gap of unknown length
* 92484: 97120: contig of 4637 bp in length
* 97121: 97220: gap of unknown length
* 97221: 104915: contig of 7695 bp in length
*
* 104916: 105015: gap of unknown length
* 105016: 112419: contig of 7404 bp in length
* 112420: 112519: gap of unknown length
* 112520: 120438: contig of 7919 bp in length
* 120439: 120538: gap of unknown length
* 120539: 128349: contig of 7811 bp in length
* 128350: 128449: gap of unknown length
* 128450: 144135: contig of 15686 bp in length
* 144136: 144235: gap of unknown length
* 144236: 155576: contig of 11341 bp in length
* 155577: 155676: gap of unknown length
* 155677: 167613: contig of 11943 bp in length
* 167620: 167719: gap of unknown length
* 167720: 180079: contig of 12360 bp in length
* 180080: 180179: gap of unknown length
* 180180: 196140: contig of 15961 bp in length.
*
FEATURES
    Location/Qualifiers
        source
            1..196140
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /strain="C57BL6/J"
                /db_xref="taxon:10090"
                /chromosome="12"
                /clone="RP23-343A21"
                /clone_lib="RPCI mouse BAC library 23"
            1..1293
                /note="assembly_fragment"
            1294..1393
                /estimated_length=unknown
            1394..2711
                /note="assembly_fragment"
            2712..2811
                /estimated_length=unknown
            2812..3945
                /note="assembly_fragment"
            3946..4045
                /estimated_length=unknown
            4046..5332
                /note="assembly_fragment"
            5333..5432
                /estimated_length=unknown
            5433..7745
                /note="assembly_fragment"
            7746..7845
                /estimated_length=unknown
            7846..10021
                /note="assembly_fragment"
            clone_end:SP6
            vector_side:right
            10022..10121
                /estimated_length=unknown
            10122..13447
                /note="assembly_fragment"
            13448..13547
                /estimated_length=unknown
            13548..16926
                /note="assembly_fragment"
            16927..17026
                /estimated_length=unknown
            17027..21840
                /note="assembly_fragment"
            21841..21940
                /estimated_length=unknown
            21941..26810
                /note="assembly_fragment"
            26811..26910
                /estimated_length=unknown
            26911..31227
                /note="assembly_fragment"
            31228..31327
                /estimated_length=unknown
            31328..36553
                /note="assembly_fragment"

```


Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

TITLE JOURNAL

Submitted (16-DEC-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Dec 16, 2005 this sequence version replaced gi:71480324.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L24694

Center clone name: 277_H_8

FEATURES

Source

Location/Qualifiers

1..154779

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="3"

/map="3"

/clone="RP24-277H8"

/clone_lib="RPCI-24 Male Mouse BAC"

1620..1660

/rpt_family="A)n"

1802..1836

/rpt_family="polypurine"

3974..4012

/rpt_family="(TG)n"

4136..4162

/rpt_family="AT-rich"

4837..4900

/rpt_family="GA-rich"

5168..5192

/rpt_family="(CAAAG)n"

5202..5227

/rpt_family="(CCAA)n"

5952..5997

/rpt_family="(CA)n"

6434..6502

/rpt_family="(GA)n"

6874..6894

/rpt_family="AT-rich"

complement(7810..7926)

/rpt_family="B1_MM"

8022..8080

/rpt_family="(TC)n"

complement(8101..8209)

/rpt_family="B1_MM"

8756..8803

/rpt_family="(TG)n"

8803..8846

/rpt_family="(GA)n"

9034..9338

/rpt_family="GA-rich"

9593..9627

/rpt_family="polypurine"

9963..10021

/rpt_family="(TG)n"

10467..10524

/rpt_family="B1F"

10660..10735

/rpt_family="ID4"

10955..11048

/rpt_family="PB1D7"

```

Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 212071)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,I., Boukhgalter,B., Camarata,J., Chang,J.,
Choepeil,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 212071)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,I., Boukhgalter,B., Camarata,J., Chang,J.,
Choepeil,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-SEP-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Sep 16, 2005 this sequence version replaced gi:28951332.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L19190
Center clone name: 176_K_16
-----
Some of the sequence contained within base pairs 104864 to the end

```

```

of the clone was stolen from accession AC121861.
FEATURES
    source
    1. .212071
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="1"
    /map="1"
    /clone="RP23-176K16"
    /clone_lib="RPCI-23 Female Mouse BAC"
    1720. .1766
    /rpt_family="(TG)n"
    repeat_region
    1949. .2094
    /rpt_family="B1_Mus2"
    repeat_region
    2101. .2138
    /rpt_family="(CAAAC)n"
    repeat_region
    complement(7429. .7512)
    /rpt_family="B3A"
    9625. .9653
    /rpt_family="(TTTG)n"
    repeat_region
    10718. .11779
    /rpt_family="Lx2"
    repeat_region
    11780. .11986
    /rpt_family="Lx2"
    repeat_region
    13111. .13142
    /rpt_family="(CA)n"
    repeat_region
    14477. .14497
    /rpt_family="(T)n"
    repeat_region
    14665. .14760
    /rpt_family="L2"
    repeat_region
    15019. .15229
    /rpt_family="B3"
    repeat_region
    15257. .15431
    /rpt_family="B2_Mm2"
    repeat_region
    15642. .15662
    /rpt_family="AT_rich"
    repeat_region
    15888. .15931
    /rpt_family="(CTCA)n"
    repeat_region
    17432. .17564
    /rpt_family="B1_Mus1"
    repeat_region
    17620. .17695
    /rpt_family="B3A"
    repeat_region
    complement(20050. .20206)
    /rpt_family="B3"
    21499. .21536
    /rpt_family="(TC)n"
    repeat_region
    21537. .21571
    /rpt_family="(CA)n"
    repeat_region
    22607. .22642
    /rpt_family="A-rich"
    repeat_region
    24493. .24520
    /rpt_family="AT_rich"
    repeat_region
    26140. .26313
    /rpt_family="B3"
    repeat_region
    26457. .26477
    /rpt_family="AT_rich"
    repeat_region
    26772. .26858
    /rpt_family="RSINE1"
    repeat_region
    27249. .27296
    /rpt_family="(TAAA)n"
    repeat_region
    29068. .30649
    /rpt_family="L1VL1"
    repeat_region
    30640. .31545
    /rpt_family="L1_Mm"
    repeat_region
    31549. .31940
    /rpt_family="L1_Mus3"
    repeat_region
    33489. .33509
    /rpt_family="AT_rich"
    repeat_region
    complement(34372. .34508)
    /rpt_family="B1_Mm"
    repeat_region
    complement(34528. .34616)
    /rpt_family="B3"
    35284. .35307

```



```

/rpt_family="CAAA)n"
35681..35831
/rpt_family="B3A"
35865..35891
/rpt_family="TATG)n"
36060..36207
/rpt_family="B4"
37216..37237
/rpt_family="AT_rich"
38584..38609
/rpt_family="AT_rich"
40954..40988
/rpt_family="AT_rich"
42272..42309
/rpt_family="AT_rich"
complement(45095..45268)
/rpt_family="B3"
45715..46090
/rpt_family="L1MB4"
46355..46512
/rpt_family="RSINE1"
46565..46954
/rpt_family="MTB"
47130..47159
/rpt_family="AT_rich"
complement(47572..47775)

Query Match          14.0%; Score 42; DB 6; Length 212071;
Best Local Similarity 57.7%; Pred. No. 0.012;
Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 24 GTGACTACGATGAGACTTGAATATCTTCATTGTAATAACAGCTAACCTCCAAATT 83
Db 50386 GTGATTTTAAATATCTAATGAATATCTTAAGTTATATAAAGCATGGTTTTTTATT 50327

Qy 84 GTACTTAAGTGGGGAGTTTATCATGAGAATGTTTAAAGCTTAATTTTCAATAAGA 143
Db 50326 ATAGTTAAACAGTGAATTTTAAAGGAGTAAATCTTAAAGGGCTAGTTACTCTGAAGT 50267

Qy 144 AGTAACCAAGA 153
Db 50266 AGTAGTAAGA 50257

RESULT 14
AC123752          223110 bp      DNA      linear      ROD 04-JAN-2005
LOCUS             Mus musculus chromosome 1, clone RP24-194N4, complete sequence.
DEFINITION        AC123752
ACCESSION          AC123752
VERSION            AC123752.8 GI:57012485
KEYWORDS           HTG.
SOURCE             Mus musculus
ORGANISM           Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 223110)
Biren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP24-194N4
Unpublished

REFERENCE
2 (bases 1 to 223110)
Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glnde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J.,

```

Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stram, K., Travers, M., Travis, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zember, L., and Zody, M.

TITLE JOURNAL REFERENCE AUTHORS

Direct Submission
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 223110)
Biren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, J., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zember, L., and Zody, M.

TITLE JOURNAL REFERENCE AUTHORS

Direct Submission
Submitted (25-NOV-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 223110)
Biren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, J., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zember, L., and Zody, M.

TITLE JOURNAL COMMENT

Direct Submission
Submitted (04-JAN-2005) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 4, 2005 this sequence version replaced gi:56089381.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information

RESULT 16

AC158640/c

LOCUS

AC158640 168914 bp DNA linear HTG 07-MAY-2005
Mus musculus clone RP23-46715, WORKING DRAFT SEQUENCE, 3 unordered
pieces.

ACCESSION

AC158640 GI:63094000

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 168914)

REFERENCE

AUTHORS

Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsebrook, S., Archer, P.,
Aredondo, H., Bandaranaike, D., Bangua, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, J., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Detamo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Druida, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernandez, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlein, K.,
Hall, B., Hanid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B.,
Howell, L. T., Hui, K. S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L. T., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
Leal, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorensu, L., Lozano, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
Nott, A., Nwako, O., Okwunonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Usmani, K., Vargo, C., Verdusco, D., Villanasa, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.

TITLE

Direct Submission

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 168914)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (07-MAY-2005)

Human Genome Sequencing Center, Baylor

College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 168914)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (07-MAY-2005)

Human Genome Sequencing Center, Baylor

----- Project Information

Center project name: MCGO

Center clone name: RP23-46715

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 168381 bases at least Q40

Consensus quality: 168573 bases at least Q30

Consensus quality: 168629 bases at least Q20

Estimated insert size: 173391; sum-of-contigs estimation

Estimated insert size: 168735; agarose-fp estimation

Quality coverage: 8x in Q20 bases; agarose-fp estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 22912: contig of 22912 bp in length

* 22913 23012: gap of unknown length

* 23013 63235: contig of 40223 bp in length

* 63236 63335: gap of unknown length

* 63336 168914: contig of 105579 bp in length.

FEATURES

source

1. 168914

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="RP23-46715"

22913..23012

/estimated_length=unknown

63236..63335

/estimated_length=unknown

ORIGIN

Query Match 13.9%; Score 41.8; DB 12; Length 168914;
Best Local Similarity 58.4%; Pred. No. 0.014;
Matches 73; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 27 ACTACGATGAAGACTTGGAAATATTCTTCATTGAAATAAAGCAGCTAACTCCCAAATTGTA 86
Db 80739 AGTACTATGAGGAGGAATATTTGTCATTAAAAAGTATATTTTCCAGATTATA 80680

Qy 87 CCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATAGAAGT 146
Db 80679 CTGAATCCAGGATGCTGATATAAATAATAGAGAGTTTCTCTGAATGTTTAACTTAGGAAT 80620

Qy 147 AACCA 151
Db 80619 GAACA 80615

RESULT 17

AC111046

LOCUS

AC111046

DEFINITION

Mus musculus chromosome 10, clone RP23-479A12, complete sequence.

ACCESSION

AC111046

VERSION

AC111046.28

KEYWORDS

HTG

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Birren, B., Nusbaum, C. and Lander, E.

TITLE

The human genome

JOURNAL

Science

286

1080-1088

1999

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

1080-1088

TITLE
JOURNAL
REFERENCE
AUTHORS

Mus musculus chromosome 10, clone RP23-479A12
2 (bases 1 to 231134)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazarro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., FitzGerald, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McGowan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Seavery, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 231134)
Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 231134)
Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNGG
Center clone name: CH230-239L13
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 196609 bases at least Q40
Consensus quality: 201630 bases at least Q30
Consensus quality: 204680 bases at least Q20
Estimated insert size: 199959; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 221906: contig of 221906 bp in length
221907: gap of unknown length
222007 233185: contig of 11179 bp in length
233186 233285: gap of unknown length
233286 234555: contig of 1270 bp in length
234556 234656: gap of unknown length
234657 234732: contig of 2777 bp in length.

FEATURES

source
1. .237432
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clones="CH230-239L13"
631. .1493
/note="clone boundary"
clone_end:T7
site:
end_sequence:BZ107018"
675. .1447
/note="clone boundary"
clone_end:T7
site:
end_sequence:BZ107018"
50656. .52220
/note="wgs_end_extension"
clone_end:T7"
91278. .93768
/note="wgs_end_extension"
clone_end:T7"
164639. .169332
/note="wgs_end_extension"
clone_end:T7"
172237. .173770
/note="wgs_end_extension"
clone_end:T7"
177148. .178958
/note="wgs_end_extension"
clone_end:T7"
complement(220855..221735)
/note="clone boundary"
clone_end:Sp6
site:
end_sequence:BZ107021"

gap
gap
gap
ORIGIN

Query Match 13.9%; Score 41.6; DB 12; Length 237432;
Best Local Similarity 53.8%; Pred. No. 0.016; Mismatches 0; Gaps 0;
Matches 86; Conservative 0; Indels 74; Indels 0; Gaps 0;

QY 37 AGACCTTGAATATCTTCAATTTGAATAAAGACGCTAACTCCCAATTTGCTACTAGG 96
Db 73971 ATATTTAATAAATACTTCTTTTGTCTAAACCCCTCTGCTAGCACTGTGATTATAAAGT 73912
QY 97 GGAGTTTATCATGAAGAATGCTTTAAATGCTTTATTTCAATAAGAGTAAACAGAGTG 156
Db 73911 CAAGTTTAGGACACTGTAAATCTCTCAGTACAAAATATCCAAATTATTACTTTTACTTTATG 73852
QY 157 TTGCAACATGCTGTTTAATAAACCCGACAAACTTCAATCAC 196
Db 73851 TTGCTATGCTCCATAACAAACAACTTAAACAAAC 73812

RESULT 19
CR792426/c

LOCUS CR792426 242992 bp DNA linear HTG 01-FEB-2006
DEFINITION Danio rerio chromosome 3 clone DREY-23A11, WORKING DRAFT SEQUENCE, 3 unordered pieces.

ACCESSION CR792426
VERSION CR792426.13 GI:86439266
KEYWORDS HTG; HTGS PHASE1; HTGS ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 242992)
AUTHORS Smith, M.
DIRECT SUBMISSION

SUBMITTED (31-JAN-2006) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight

On Feb 2, 2006 this sequence version replaced gi:85814118.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK23A11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 242229 bases at least Q40
Consensus quality: 242331 bases at least Q30
Consensus quality: 242425 bases at least Q20
Insert size: 242792; sum-of-contigs
Insert size: 235334; 12.6% error; agarose-fp
Quality coverage: 7.24x in Q20 bases; sum-of-contigs Quality coverage: 7.83x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 112730: contig of 112730 bp in length


```

* 112731 112830: gap of 100 bp
* 112831 240283: contig of 127453 bp in length
* 240284 240383: gap of 100 bp
* 240384 242992: contig of 2609 bp in length.
FEATURES
    source
        1..242992
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="RZPD:HUKGB735A11230"
            /db_xref="taxon:7955"
            /chromosome="3"
            /clone="DKEY-23A11"
            /clone_lib="DanioKey"
            1..112730
                /note="assembly_fragment:00647"
                /fragment_chain:1
            112831..240283
                /note="assembly_fragment:02028"
                /fragment_chain:1
            240384..242992
                /note="assembly_fragment:03120.0"
ORIGIN
Query Match      13.9%; Score 41.6; DB 12; Length 242992;
Best Local Similarity 50.0%; Pred. No. 0.016;
Matches 104; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
Qy 42 TGAATATTCCTTCATTGGAATAAACAGCTAACCTCCCAAAATGTTACCTAACTAGGGAGT 101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148000 TGAAGTGTTCACAGTATTGGAATGTGTAATACACTGTATAGTCTTCATTGTGCAAT 147941
Qy 102 TTATCATGGAAGAAATGTTTAAATGCTTAAATTTTCAATAGAAAGTAAACAGAGTGTTCGA 161
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147940 TTAGTACATACAAATATCTTTTAAGCTACAATTTATTAATTTCTTATACCAAAATAGTTCCAC 147881
Qy 162 ACATGCTGTTAAATAACCCGCAAACTCAATCACTATAGCTAGTAGTAGTCAATCTG 221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147880 AATTGCTTGAATTTATCACACAGCCTTAAAGCACATGCAATGCAATGATAAACGTTCCATTG 147821
Qy 222 CAAGGATCCCAAGTAGTAAACCAATTTTG 249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147820 CAATGATCCCATATATATGAACTGTG 147793

RESULT 20
AC094947/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-6C17, WORKING DRAFT SEQUENCE, 3
ACCESSION AC094947
VERSION AC094947.6 GI:30467040
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
    1 (bases 1 to 222752)
    Muzny D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
    Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
    Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
    Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
    Biswal O., Blair J., Blankenburg K., Blyth P., Brown M.,
    Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
    Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
    Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
    Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza B.,
    Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
    Delgado O., Denson S., Detamo C., Ding Y., Dinh H., Divya K.,
    Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
    Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
    Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
    Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,

```

Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idelbird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Maheshwari, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, W.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, J., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vallas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 222752)
 Worley, K.C.
 Direct Submission
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222752)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24941592.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSUG
 Center clone name: CH230-6C17
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 213463 bases at least Q40
 Consensus quality: 215000 bases at least Q30

Consensus quality: 216043 bases at least Q20
 Estimated insert size: 22326; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: this is a "working draft" sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 219872: contig of 219872 bp in length
 * 219873 219972: gap of unknown length
 * 219973 221181: contig of 1209 bp in length
 * 221182 221281: gap of unknown length
 * 221282 222752: contig of 1471 bp in length.

FEATURES

Location/Qualifiers
 1. .222752
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clones="CH230-6C17"
 219873. .219972
 /estimated_length=unknown
 221182. .221281
 /estimated_length=unknown

ORIGIN

Query Match 13.8%; Score 41.4; DB 12; Length 222752;
 Best Local Similarity 48.1%; Pred. No. 0.019;
 Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 58 TGAATAACAGCTAACTCCAAATGTACCTACTAGGAGTTTATCATGAGAATG 117
 |||||
 16821 TGTAAATACAGATATGTTGCAAAATATATGTAATTTGTGCATATGCTGCACAGATA 16762
 |||||
 118 TTTAAATGCTTAATTTCAATAAGACGATACCAAGAGTGTGCAACATGCTGTTAAATAA 177
 |||||
 16761 AATATATCAATTTA 16702
 |||||
 178 CCGGACAAATCTCAATCACTAGCTAGTGTAGTGTGATTCGCAAGGATCCAGAGTA 237
 |||||
 16701 CAGATTTCTATAACTATCTGATCTGTTGCTATTTTAAATTTAAAGAACTACCTAATA 16642
 |||||
 238 ACCAGTATTTGGAATGCAATGTTGAACGACCATCTATATATCTGCTGATTAGGAAG 297
 |||||
 16641 ATTGAAGAATATAGATATATATATATTTTACTTTTGTATCTGTTTGTATAGGTAG 16582
 |||||
 298 TAT 300
 16581 GAT 16579

RESULT 21
 AC131839/c
 LOCUS AC131839 253693 bp DNA linear HTG 13-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-213, *** SEQUENCING IN PROGRESS ***,
 2 unordered pieces.
 AC131839
 VERSION AC131839.3 GI:24941378
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 253693)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.B., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Caron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davlaney,M.L., Davis,C., Davy-Cartoll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
 Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Ogunu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sison,I.,
 Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 253693)
 Worley,K.C.
 Direct Submission
 Submitted (26-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 253693)
 Worley,K.C.
 Direct Submission
 Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 13, 2002 this sequence version replaced gi:23664809.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequences contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: TUQL
 Center clone name: CH230-213
 ----- Summary Statistics

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: TUCU
 Center clone name: CH230-2C7
 ----- Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 254993 bases at least Q40
 Consensus quality: 259665 bases at least Q30
 Consensus quality: 262708 bases at least Q20
 Estimated insert size: 273866; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 5545: contig of 5545 bp in length
 * 5546 5645: gap of unknown length
 * 5646 10296: contig of 4651 bp in length
 * 10297 10396: gap of unknown length
 * 10397 258638: contig of 248242 bp in length
 * 258639 258738: gap of unknown length
 * 258739 260235: contig of 1497 bp in length
 * 260236 260335: gap of unknown length
 * 260336 261450: contig of 1115 bp in length
 * 261451 261550: gap of unknown length
 * 261551 262936: contig of 1386 bp in length
 * 262937 263036: gap of unknown length
 * 263037 263439: contig of 1313 bp in length
 * 264450 264449: gap of unknown length
 * 264450 265703: contig of 1254 bp in length
 * 265704 265803: gap of unknown length
 * 265804 267592: contig of 1789 bp in length
 * 267593 267692: gap of unknown length
 * 267693 269812: contig of 2120 bp in length
 * 269813 271927: gap of unknown length
 * 269913 271927: contig of 2015 bp in length.

FEATURES

source

1. 271927
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-2C7"
 5546..5645
 /estimated_length=unknown
 8661..10296
 /note="wgs contig"
 10297..10396
 /estimated_length=unknown
 258639..258738
 /estimated_length=unknown
 260236..260335
 /estimated_length=unknown

gap

261451..261550
 /estimated_length=unknown
 262937..263036
 /estimated_length=unknown
 264350..264449
 /estimated_length=unknown
 265704..265803
 /estimated_length=unknown
 267593..267692
 /estimated_length=unknown
 269813..269912
 /estimated_length=unknown

ORIGIN

Query Match 13.8%; Score 41.4; DB 12; Length 271927;
 Best Local Similarity 48.1%; Pred. No. 0.019;
 Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 58 TGAATAAACAGCTTAACCTCCCAAAATTGACCTAACTAGGGAGTTTATCATGAAGAAATG 117
 |||||
 Db 49459 TGTAAACAGATATTGTTGCAAAATTTATGTAAATTGTGCCATATGCTGCACAGATA 49518
 |||||
 QY 118 TTTAAATGCTTAATTTTCAAAATAGAACTAACCAAGAGTTTGCAACATGCTGCTTAAATAA 177
 |||||
 Db 49519 AATATATCAATTAATGTTATATATATATATACATCATGTATACATAATGCATATATATAA 49578
 |||||
 QY 178 CCCGACAACTTCAATCACTATAGCTGTAGTAGTGCAATCTGCAAGGATCCACAGATA 237
 |||||
 Db 49579 CAGATTTCTATACTATCTGATCTGTTGCTATTATTTAAATTAAGAACTACCTAATA 49638
 |||||
 QY 238 ACCAGTATTTTGGAAATGCAATGTTGAACCGACCACTACTAATATCTGCTGATAGGAAG 297
 |||||
 Db 49639 ATTGAAGAAATTAGATATTATATTTTACCTTTTATGTTTGTGTTTGTATTAGTAGT 49698
 |||||
 QY 298 TAT 300
 |||||
 Db 49699 GAT 49701

RESULT 23

AP004980/c AP004980 23709 bp DNA linear PLN 22-JUL-2003
 LOCUS Lotuscorniculatus var. japonicus genomic DNA, chromosome 3,
 DEFINITION clone:Lfj13N17, TM0436a, complete sequence.
 ACCESSION AP004980
 VERSION AP004980.1 GI:21907998
 KEYWORDS HTG.
 SOURCE Lotuscorniculatus var. japonicus (Lotus japonicus)
 ORGANISM Lotuscorniculatus var. japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
 Lotus.
 REFERENCE 1
 AUTHORS Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.
 TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence
 Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
 Regions of the Genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 23709)
 AUTHORS Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research; 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: ynk@kzusa.or.jp,
 URL: <http://www.kazusa.or.jp>, Tel:81-438-52-3935,
 Fax:81-438-52-3934)

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 23709
 /location/Qualifiers
 /organism="Lotus corniculatus var. japonicus"
 /mol_type="genomic DNA"
 /variety="japonicus"
 /db_xref="taxon:34305"

```

/chrnosome="3"
/clone="Lj113N17"
/clone_lib="Lj1 library"
/notes="TM0436a, a part of TAC clone:TM0436.
synonym: Lotus japonicus"

ORIGIN
Query Match      13.6%; Score 40.8; DB 4; Length 23709;
Best Local Similarity 49.5%; Pred. No. 0.031;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

4 ATTGACAGGCTAAATCTAGTACGACGAGACTTGAATATATCTTCATTGGAAT 63
|||||
10080 ATTGGTTATCTAAATGAGAATTTCAAAAAAAGAGACATACATCTCAATCTTGCT 10021
|||||

64 AAACAGCTAACTCCCAATTTGTACTAACTAGGGAGCTTTATCATGAGAAATGTTTAAA 123
|||||
10020 AAAATATTTTCCCATTTTACAAAATTAGGAAAATAAAAAAATAAAAAATCTAT 9961
|||||

124 TGCTTAATTTCAAAATGAAGTAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCCGAC 183
|||||
9960 TATATAAATCAAGAGAAGACTCAACAAAAAGGTGATTTCTTCCATTCATGAC 9901
|||||

184 AACTTCAATCACTATAGCTCTAGTAGTGC 215
|||||
9900 AATGCTCATGTTGCTGTGTGCTGTGAATGC 9869
|||||

RESULT 24
BX255970      158755 bp      DNA      linear      VRT 17-APR-2005
LOCUS
DEFINITION
  Zebrafish DNA sequence from clone CH211-136G2 in linkage group 4
  Contains the gene for a novel slow skeletal troponin T family
  protein, and a CpG island, complete sequence.
ACCESSION
  BX255970
VERSION
  BX255970.8 GI:49203508
KEYWORDS
  HTG.
SOURCE
  Danio rerio (zebrafish)
  Danio rerio
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
  Cypriniformes; Cyprinidae; Danio.
  1 (bases 1 to 158755)
  Dunn.M.
  Direct Submission
  Submitted (02-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  zfish-help@sanger.ac.uk Clone requests:
  http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
  On Jun 24, 2004 this sequence version replaced gi:38228952.
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: zfish-help@sanger.ac.uk
  -----
  The following abbreviations are used to associate primary accession
  numbers given in the feature table with their source databases:
  Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
  on the WormPEP database can be found at
  http://www.sanger.ac.uk/Projects/C_elegans/wormpep
  Zebrafish pUC subclones occasionally display inconsistency over the
  length of mononucleotide A/T runs and conserved TA repeats. Where
  this is found the longest good quality representation will be
  submitted.
  CH211-136G2 is from a CHORI-211 BAC library
  VECTOR: pTARBAC2.1
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one subclone; and the assembly was confirmed by restriction digest,
  except on the rare occasion of the clone being a YAC.
  Any regions longer than 1kb tagged as misc-feature 'unsure' are
  part of a tandem repeat of more than 10kb in length where it has
  not been possible to anchor the base differences between repeat
  copies. The region has been built up based on the repeat element
  to match the total size of repeat indicated by restriction digest,
  but repeat copies may not be in the correct order and the usual
  finishing criteria may not apply.
  Location/Qualifiers
    source
      1..158755
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /chromosomes="4"
        /clone="CH211-136G2"
        /clone_lib="CHORI-211"
    misc_feature
      1
        /note="Clone left end: CH211-136G2"
    gene
      join(836..882,4803..4853,7092..7096,11651..11672,
      11766..11823,14755..14871,15177..15254,15358..15471,
      15805..15917,16987..17077,18279..18319,24154..24209)
        /locus_tag="CH211-136G2.1-002"
    mRNA
      join(836..882,4803..4853,7092..7096,11651..11672,
      11766..11823,14755..14871,15177..15254,15358..15471,
      15805..15917,16987..17077,18279..18319,24154..24209)
        /locus_tag="CH211-136G2.1-002"
        /product="novel slow skeletal troponin T family protein"
        /note="match: ESTs: Em:CF347331.1 Em:CF348408.1
        Em:CK142922.1"
    gene
      join(840..882,4803..4853,7092..7096,11079..11105,
      11651..11672,11766..11823,14755..14871,15177..15254,
      15358..15471,15805..15917,16987..17077,18279..18312)
        /locus_tag="CH211-136G2.1-004"
    mRNA
      join(840..882,4803..4853,7092..7096,11079..11105,
      11651..11672,11766..11823,14755..14871,15177..15254,
      15358..15471,15805..15917,16987..17077,18279..18312)
        /locus_tag="CH211-136G2.1-004"
        /product="novel slow skeletal troponin T family protein"
        /note="match: ESTs: Em:CNS04044.1"
    gene
      join(848..882,4803..4853,7092..7096,11651..11672,
      11766..11823,14755..14871,15177..15254,15358..15471,
      15805..15917,16987..17077,18517..18554,24154..24411)
        /locus_tag="CH211-136G2.1-001"
    mRNA
      join(848..882,4803..4853,7092..7096,11651..11672,
      11766..11823,14755..14871,15177..15254,15358..15471,
      15805..15917,16987..17077,18517..18554,24154..24411)
        /locus_tag="CH211-136G2.1-001"
        /product="novel slow skeletal troponin T family protein"
        /note="match: ESTs: Em:CF266249.1 Em:CK029104.1"
    gene
      join(4806..4853,7092..7096,8006..8026,9552..9557,
      9805..9825,9977..9997,10413..10430,10527..10550,
      11079..11105,11651..11672,11766..11823,14755..14871,
      15177..15254,15358..15471,15805..15917,16987..17077,
      18517..18554,24154..24199)
        /locus_tag="CH211-136G2.1-003"
    mRNA
      join(4806..4853,7092..7096,8006..8026,9552..9557,
      9805..9825,9977..9997,10413..10430,10527..10550,
      11079..11105,11651..11672,11766..11823,14755..14871,
      15177..15254,15358..15471,15805..15917,16987..17077,
      18517..18554,24154..24199)
        /locus_tag="CH211-136G2.1-003"
        /product="novel slow skeletal troponin T family protein"
        /note="match: ESTs: Em:CNS08154.1"
    CDS
      join(4816..4853,7092..7096,8006..8026,9552..9557,
      9805..9825,9977..9997,10413..10430,10527..10550,
      11079..11105,11651..11672,11766..11823,14755..14871,
      15177..15254,15358..15471,15805..15917,16987..17077,
      18517..18554,24154..24199)
        /locus_tag="CH211-136G2.1-003"
        /standard_name="OTTDARP0000007299"
        /note="match: proteins: Tr:Ol3096"
        /codon_start=1

```

```

/locus_tag="CH211-136G2.3-001"
join(31521..31816,34657..34755,34931..35208)
/locus_tag="CH211-136G2.3-001"
/product="novel transcript"
/note="match: ESTs: Em:CN174698.1"
join(36670..36967,38896..38994,39172..39449)
/locus_tag="CH211-136G2.4-001"
/note="match: ESTs: Em:CN174698.1"
join(36670..36967,38896..38994,39172..39449)
/locus_tag="CH211-136G2.4-001"
/product="novel transcript"
/note="match: ESTs: Em:CN174698.1"

Query Match      13.6%; Score 40.8; DB 11; Length 158755;
Best Local Similarity 53.0%; Pred. No. 0.031;
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 36 AAGACTTGAATATTTCTTCATTGCAATAAAGACAGTCACTCCCAATTTGTACCTAACATAG 95
Db 150328 AATTATTTGCAATATTTACACTTCAGAAAAAATACATTTGGTCACACTTGTACATTTTAT 150387
QY 96 GGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATAAGAAAGTAAACAGACT 155
Db 150388 TTTACTATACATAAACTCAGTATTAATACCATATATTTTAAATAAACATTACTTAAG 150447
QY 156 GTTGCAACATGCTGTTTAAATAACCCGACAAACTTCAATCACTAT 199
Db 150448 TACAATACATTTTACTTAAAAAAAATACAAATTTTAAAAACAGTAT 150491

RESULT 25
AL356073/c
LOCUS
DEFINITION
AL356073 Human DNA sequence from clone RP11-81303 on chromosome 10, complete
sequence.
ACCESSION
AL356073 GI:16943998
VERSION
HTG.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 165917)
Pelán,S.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 15, 2001 this sequence version replaced gi:15590948.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/
This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chr10
RP11-81303 is from the library RPCI-11.3 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
-----
Center: Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least

```


one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```

FEATURES
    source
        1..165917
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="10"
            /clone="RP11-81303"
            /clone_lib="RPC1-11.3"
    misc_feature
        1 /note="Clone_left_end: RP11-81303"
    misc_feature
        46771..46804
            /note="Sequence from uni-directional dGTP big dye
            terminator reads only."
    misc_feature
        46771..46803
            /note="Sequence from clone PCR only."
    misc_feature
        46804
            /note="Gap size estimated to be approximately 330bp by
            restriction digest data."
    misc_feature
        51135
            /note="Clone_right_end: RP11-765C10"
    misc_feature
        66101
            /note="Clone_left_end: RP11-129G17"
    misc_feature
        165917
            /note="Clone_right_end: RP11-81303"
ORIGIN
Query Match      13.6%; Score 40.8; DB 5; Length 165917;
Best Local Similarity 55.7%; Pred. No. 0.031;
Matches 78; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 33 ATGAAGACTTGAAATATCTTCATTCGAAATAAACAGTAACTCCCAATTTGTACCTAAC 92
DB 55035 ATTAGGTATAGATACTCTCTATCATAGAAGATAGTACCAAAATCTACATCA 54976
QY 93 TAGGGGAGTTATCATGAAGAAATGTTTAAATGCTTAATTTCAATTAAGAAGTAAACAG 152
DB 54975 AAATGTCCTGCTTAATGCTGAATGTTAAATAATTTACATTTGATGATCAGAAAAAGGCA 54916
QY 153 AGTGTGCAACATCTCTTA 172
DB 54915 AGGATGCAAGCTTGCTTTCA 54896

RESULT 26
AC135483/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-261E14, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
ACCESSION AC135483
VERSION AC135483.2 GI:25139065
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 166494)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Blawalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

```

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Johnson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhea,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoeleneh,O., Okwuonu,G., Olarnpunsegon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Primus,E., Pu,L.-L., Plopper,F., Poindexter,A., Popovic,D., Reeves,K., Regier,M.A., Reigh,R., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Rejos,Z., Usmani,K., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,K., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 166494)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (17-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 166494)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:24080623.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBWY
Center clone name: CH230-261E14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 151587 bases at least Q40

Consensus quality: 154840 bases at least Q30
 Consensus quality: 156515 bases at least Q20
 Estimated insert size: 155596; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 3 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* 1 145399: contig of 145399 bp in length
 * 145400: gap of unknown length
 * 145500: contig of 19262 bp in length
 * 164761: gap of unknown length
 * 164762: gap of unknown length
 * 164862: contig of 1633 bp in length.

FEATURES

source
 1..166494

/organism="Rattus norvegicus"
 /mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-261B14"

3174..3995

/note="clone boundary"

clone_end:Sp6

site:

end_sequence:BZ245159"

72113..72978

/note="clone boundary"

clone_end:T7

site:

end_sequence:BZ245158"

75147..76243

/note="wgs_end_extension"

clone_end:T7"

76585..80486

/note="wgs_end_extension"

clone_end:T7"

145400..145499

/estimated_length=unknown

145500..147557

/note="wgs_end_extension"

clone_end:T7"

164762..164861

/estimated_length=unknown

ORIGIN

Query Match 13.5%; Score 40.4; DB 12; Length 166494;
 Best Local Similarity 48.7%; Pred No. 0.042;
 Matches 110; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
 22 AAGTGACTACGATGAGAGCTTGAATATTTCTTCATTGGAATATAACAGCTAATCCCAAA 81
 DB |||||
 122057 AACTGAGAAATATATAAAATCAAAAGTTGTGGAATGTCATAGGAATACAGAGGCAAG 121998
 QY |||||
 82 TTGTACCTAAGTGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTTCAAATAA 141
 DB |||||
 121997 TTTTATCAAAAGAAATPAAGACAGATGGAAGAGAAATCTCAGGCTCAGAATATACAAAAGA 121938
 QY |||||
 142 GAAGTAACGAGTGTCTCAACATCTGTTAAATAACCGCAAACTTCAATCACTATAG 201
 DB |||||
 121937 AAAGTGATATATTTGGTCAAGAAATATGTTGAATATGACAAAACCTTCTAGCACAAG 121878
 QY |||||
 202 CTGTAGTAGAGTGCATTCTCGAAGAGATCCCAAGTAACCAAGTATTT 247
 DB |||||
 121877 ATCCAGAAATCCAAACCTATATAAAGACCAAACTTACAAATAATT 121832

RESULT 27

AC142454/c

LOCUS

DEFINITION

AC142454

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

AC142454 183223 bp DNA linear ROD 30-JUL-2005
 Mus musculus BAC clone RP23-313G1 from chromosome 3, complete.
 sequence.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC Library has been constructed by Kazutoyo Oseawa and
 Minako Tateno in the laboratory of Pieter de Jong
 (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality
 >=30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence. The
 assembly was confirmed by restriction digest.
 This finishing standard has slightly changed from the previous
 Human standard. Specifically, standards for regions of low sequence
 complexity (such as dinucleotide repeats and small unit tandem
 repeats) have been relaxed.

brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

```

FEATURES             source
    source
        1..183223
            Location/Qualifiers
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /chromosome="3"
                /clone="RP23-313G1"
                /clone_lib="RPCI-23"
                /clone_114411
            misc_feature
                114260..114411
            /note="Sequence derived from PCR product of project DNA"

ORIGIN
Query Match      13.5%; Score 40.4; DB 6; Length 183223;
Best Local Similarity 52.4%; Pred. No. 0.042;
Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY      60 AATATAACAGCTACTCCCAATGTACTAACTAGGAGGAGTTTATCATGAAGAAATGTT 119
DB      7285 AACTGAACACACTCTTCTCAATGATACCTTGGTCAAGGAGGATTAAGAAAAATTGA 7226

QY      120 TAAATGCTTAATTTCAATAAGAAAGTACAGAGTGTGCAACATCTGCTTTAATAAC 179
DB      7225 GGACTTTTATGAGTTCAATGAATAAGAACACAGCATACCCCAATTTATTGACATAATG 7166

QY      180 CGACAACACTCAATCACTATAGCTGTAGTAGAGTGCATTTCTGCAGGATC 229
DB      7165 AAAGGATTTTAAAGAGAAACTCATAGCTGAGTGCCTCCAAAAAGAAAC 7116

RESULT 28
AC139360/c
LOCUS      AC139360      186836 bp      DNA      linear      ROD 20-FEB-2004
DEFINITION Mus musculus chromosome 3, clone RP23-324F4, complete sequence.
ACCESSION      AC139360
VERSION      AC139360.8      GI:42716198
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 186836)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 3, clone RP23-324F4
Unpublished

REFERENCE
2 (bases 1 to 186836)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Maddonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

FEATURES             source
    source
        1..186836
            Location/Qualifiers
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /chromosome="3"
                /map="3"
                /clones="RP23-324F4"
                /clone_lib="RPCI-23 Female Mouse BAC"
                /clone_1146
            misc_feature
                1..6
                /note="clone_boundary"

```

REFERENCE AUTHORS

Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186836)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Maddonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 186836)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Maddonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (20-FEB-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 20, 2004 this sequence version replaced gi:41058797.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19543

Center clone name: 324_F_4

FEATURES source

1..186836
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/clones="RP23-324F4"
/clone_lib="RPCI-23 Female Mouse BAC"
1..6
/note="clone_boundary"

misc_feature

TITLE

JOURNAL

JOURNAL

Submitted (08-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Feb 21, 2003 this sequence version replaced gi:27819585.

COMMENT

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: M_BA0187E03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseagawa and Minako Tatenoi in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124440.

FEATURES

source

Location/Qualifiers

1..192918
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="3"

/clone="RP23-187B3"
 /clone_lib="RPCI-23"

784..1161

/rpt_family="ERVK"

1827..2046

/rpt_family="B4"

2408..3581

/rpt_family="L1"

4262..4541

/rpt_family="MaLR"

4545..4933

/rpt_family="MaLR"

7970..8111

/rpt_family="L2"

9273..9466

/rpt_family="B4"

10761..10836

/rpt_family="L1"

10809..10940

/rpt_family="L1"

12991..13202

/rpt_family="B2"

25930..26260

/rpt_family="L1"

repeat_region
 29078..29212
 /rpt_family="B4"
 repeat_region
 30956..31120
 /rpt_family="RMER15"
 repeat_region
 32294..32449
 /rpt_family="B2"
 repeat_region
 32937..32989
 /rpt_family="ERV1"
 repeat_region
 33851..33968
 /rpt_family="B4"
 repeat_region
 34965..35371
 /rpt_family="MaLR"
 repeat_region
 35375..36428
 /rpt_family="MaLR"
 repeat_region
 36429..36737
 /rpt_family="MaLR"
 repeat_region
 38654..38794
 /rpt_family="Alu"
 repeat_region
 39154..39237
 /rpt_family="Alu"
 repeat_region
 41191..41297
 /rpt_family="Alu"
 repeat_region
 51184..51569
 /rpt_family="MaLR"
 repeat_region
 51730..51952
 /rpt_family="L1"
 repeat_region
 52503..52672
 /rpt_family="B4"
 repeat_region
 52775..52872
 /rpt_family="MIR"
 repeat_region
 53927..54116
 /rpt_family="B2"
 repeat_region
 54118..54230
 /rpt_family="B4"
 repeat_region
 63831..64645
 /rpt_family="RMER1B"
 repeat_region
 64676..64821
 /rpt_family="L1"
 repeat_region
 64899..64946
 /rpt_family="ID"
 repeat_region
 65886..66372
 /rpt_family="ERVK"
 repeat_region
 66375..66543
 /rpt_family="ERVK"
 repeat_region
 66549..66642
 /rpt_family="ERVK"
 repeat_region
 66861..66968
 /rpt_family="ERVK"
 repeat_region
 67443..67522
 /rpt_family="ERVK"
 repeat_region
 67853..67913
 /rpt_family="ERVK"
 repeat_region
 67952..68163
 /rpt_family="ERVK"
 repeat_region
 68202..68244
 /rpt_family="ERVK"
 repeat_region
 68247..68429
 /rpt_family="ERVK"
 repeat_region
 68432..68918
 /rpt_family="ERVK"
 repeat_region
 68921..69089
 /rpt_family="ERVK"
 repeat_region
 69095..69188
 /rpt_family="ERVK"
 repeat_region
 69193..69270
 /rpt_family="L1"
 repeat_region
 69722..69861
 /rpt_family="B4"
 repeat_region
 72101..72229
 /rpt_family="ERV1"
 repeat_region
 72279..72534
 /rpt_family="ERV1"
 repeat_region
 72540..72588

[illegible]

* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will be preserved.

* 1 208908: contig of 208908 bp in length
 * 208909 209008: gap of unknown length
 * 209009 210058: contig of 1050 bp in length
 * 210059 210158: gap of unknown length
 * 210159 211674: contig of 1516 bp in length
 * 211675 211774: gap of unknown length
 * 211775 213458: contig of 1684 bp in length.

FEATURES

source
 1. 213458
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clones="CH230-514P9"
 misc_feature
 1. 1230
 /note="wgs contig"
 misc_feature
 2053. 3154
 /note="wgs contig"
 gap
 208909..209008
 /estimated_length=unknown
 gap
 210059..210158
 /estimated_length=unknown
 gap
 211675..211774
 /estimated_length=unknown

ORIGIN

Query Match 13.5%; Score 40.4; DB 12; Length 213458;
 Best Local Similarity 48.7%; Pred. No. 0.041;
 Matches 110; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
 QY 22 AAGTGACTGACGACGACTGTAATATCTTCATTGGAATAACACGCTAACCTCCAAA 81
 Db 126364 AACTGGAATAATTATAAAATCAAAATGTTGGAAATGTCTAGGAACTCAGAGGCAAG 126423
 QY 82 TTGTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATCTTAAATTTTCAAAATAA 141
 Db 126424 TTTTATCAAGAATAAGACAGATGGAAGAAATCTCAGGCTCAGATATACAAAGA 126483
 QY 142 GAAGTAACGAGAGTGTGCAACATGCTGTTTAAATAACCCGCAAACTTCAATCACTATAG 201
 Db 126484 AAAGTGAATATATTGGTCAAGAAATGTTGAAATTTGAACAAAACCTTCTAGCACAAAAG 126543
 QY 202 CTGTAGTAGAGTGATCTGCAAGGATCCGACGAGTAACCGATTTT 247
 Db 126544 ATCCAGAAATCCAAACCTATATAAAGACCAACCTACAATAATT 126589

RESULT 31

AC095434/c
 LOCUS AC095434 222392 bp DNA linear HTG 09-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-6P20, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AC095434
 VERSION AC095434.9 GI:30467849
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Rattus.
 1 (bases 1 to 222392)
 Muzny D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
 Allen C., Allen H., Alsbrooks S., Amin A., Anguiano D.,
 Anyalebechi V., Ayodeji A., Ayodeji M., Baca E., Baden H.,
 Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
 Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
 Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
 Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
 Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,

Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
 Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
 Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
 Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
 Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
 Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
 Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
 Gebregorgis E., Geer K., Gill R., Grady W., Guerra W., Guevara W.,
 Gunaratne P., Haaland W., Hamill C., Hamilton C., Hamilton K.,
 Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
 Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
 Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
 Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
 Karpathy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
 Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
 Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
 Lorensuhewa L., Loulseghe H., Lozado R.J., Lu X., Ma J.,
 Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,
 Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
 Mawhinney S., McLeod M.P., McNeill T.Z., Meenen E.,
 Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
 Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
 Nankervis C., Neal D., Newton N., Nguyen N., Norris S.,
 Nwaokemele O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
 Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C.,
 Plopper F., Poindexter A., Popovic D., Primus E., Pu L.-L.,
 Puazo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
 Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F.,
 Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
 Sanders W., Savery G., Scherer S., Scott G., Shatsman S., Shen H.,
 Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajd D.,
 Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
 Steinle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
 Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K.,
 Valas R., Vera V., Villasana D., Waldron L., Walker B., Wang J.,
 Wang Q., Wang S., Warren R., Wlezyk R., Wooden H., Worley K.,
 Williams G., Willson R., Wlezyk R., Wlezyk R., Worley K.,
 Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,
 Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von
 Niederhausen A., Weiss R., Smith D.R., Holt R.A., Smith H.O.,
 Weinstock G. and Gibbs R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 222392)
 Worley K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222392)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24941161.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

```

----- Project Information
Center project name: GCGI
Center clone name: CH230-6P20
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 189423 bases at least Q40
Consensus quality: 192508 bases at least Q30
Consensus quality: 194817 bases at least Q20
Estimated insert size: 199567; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
  1 222392: contig of 222392 bp in length.
  Location/Qualifiers
    1..222392
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-6P20"
    misc_feature
      1..1189
        /note="wgs_contig"
    misc_feature
      211317..213363
        /note="wgs_contig"
ORIGIN
Query Match      13.5%; Score 40.4; DB 12; Length 222392;
Best Local Similarity 47.9%; Pred. No. 0.041;
Matches 116; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

59  GAAATAAACAGCTAACTCCCAATGTGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
DB  23487  GTAATACAGATATGTTGCAAAATTTATGTATTAATTTGCCATATGCTGCACAGATAA 23428

119  TTAATGCTTAATTTTCAAAATGAAGATPACAGAGTGTGCAACATGCTGTTAAATAAC 178
DB  23427  ATATATCATTTAATGTTTATAATATATATACATACATGATATCAATATGATATATAAC 23368

179  CCGACAACTTCAATCATTAGCTGTAGTAGTGCAATTCGCAAGATCCAGAGTAA 238
DB  23367  AGATTTCTATAACCTATCTGATCTGTTGTTCTATTATTTAAAGAACTACCTAATAA 23308

239  CCAGTATTTTGGAAATGCAATGTTGCAACCGACCTACTAATTTATCTGCTGATTAGGAAGT 298
DB  23307  TTGAAGATTTATAGTATATATAATTTTACTTTTGTAGTTGTTGATTGATTAGGTAGG 23248

299  AT 300
DB  23247  AT 23246

Sequence split into 28 fragments LOCUS AE015927 Accession AE015927
Fragment Name      Begin      End
AE015927_00        1      110000
AE015927_01      100001      210000
AE015927_02      200001      310000
AE015927_03      300001      410000
AE015927_04      400001      510000
AE015927_05      500001      610000
AE015927_06      600001      710000
AE015927_07      700001      810000

```

```

AE015927_08      800001      910000
AE015927_09      900001     1010000
AE015927_10     1000001     1110000
AE015927_11     1100001     1210000
AE015927_12     1200001     1310000
AE015927_13     1300001     1410000
AE015927_14     1400001     1510000
AE015927_15     1500001     1610000
AE015927_16     1600001     1710000
AE015927_17     1700001     1810000
AE015927_18     1800001     1910000
AE015927_19     1900001     2010000
AE015927_20     2000001     2110000
AE015927_21     2100001     2210000
AE015927_22     2200001     2310000
AE015927_23     2300001     2410000
AE015927_24     2400001     2510000
AE015927_25     2500001     2610000
AE015927_26     2600001     2710000
AE015927_27     2700001     279251
Continuation (6 of 28) of AE015927 from base 500001 (AE015927 Clostridium tetani E88, cc

Query Match      13.3%; Score 40; DB 15; Length 110000;
Best Local Similarity 51.1%; Pred. No. 0.057;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY  37  AGACTTGAATATTTCTTCATTTGAAATAAAGAGCTAACTCCCAATTTGCTACTACTAGG 96
DB  21776  AGAATTAAATAGTAGTTTATAAAAAGGGCATATTACTTTCAATATACCATATGAAA 21835

QY  97  GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATATGAAGATTAACCGAGTG 156
DB  21836  GCAGTATGCTCTATAAAATAATATAAATGTTGAAAAATTTATATAAACACTATACCCATA 21895

QY  157  TTGCAACATGCTGTTAAATACCCGACAACTTCAATCACTATAGTCTAGTAGAGTGCA 216
DB  21896  TTATAAATGTAAATTTAAATTTGCAATCAACCTTTACTATATATATAAATAATGGGAT 21955

QY  217  TTCT 220
DB  21956  TTTT 21959

RESULT 33
AL137018/c
LOCUS AL137018 157168 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-133M9 on chromosome 9q13-21.33
            Contains the 3' end of the RORB gene for RAR-related orphan
            receptor B protein (RZRB, ROR-BETA), complete sequence.
ACCESSION AL137018
VERSION AL137018.9 GI:14141220
KEYWORDS HTG; receptor; ROR-BETA; RORB; RZRB.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE 1 (bases 1 to 157168)
AUTHORS Tracey,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT Clone requests: clonerequest@sanger.ac.uk
On May 17, 2001 this sequence version replaced gi:13751249.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9

```


RP11-133M9 is from the library RPI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

```

source
    Location/Qualifiers
    1..157168
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="9"
        /map="q13-21.33"
        /clone="RP11-133M9"
        /clone_lib="RPI-11.1"
        /note="Clone right end: RP11-171A24"
    100
misc_feature
    81385
        /note="Clone left end: RP11-64C9"
        /join(AL355674.10:140352..140970,98495..155406)
        /gene="RORB"
        /locus_tag="RP11-133M9.1-001"
        /join(AL355674.10:140352..140970,98495..98580,
        102844..102985,110627..111028,128830..128951,
        130687..130819,133701..133808,136004..136114,
        140002..140114,153676..155406)
        /gene="RORB"
        /locus_tag="RP11-133M9.1-001"
        /product="RAR-related orphan receptor B"
        /note="match: ESTs: AA084457.1 AA780836.1 AL118995.1
        AL550102.1 AL550103.1 AL691650.1 AL691651.1 AL711967.1
        AL712330.1 AL444764.1 BF204507.1 BG165895.1 BM553788.1
        BM669680.1 BM698896.1 BU169404.1 CA397109.1 T26966.1
        match: cDNAs: BC008831.1 L14611.1 U04897.1 U04898.1
        U04899.1 Y08638.1 Y08639.1"
        /join(AL355674.10:140964..140970,98495..98580,
        102844..102985,110627..111028,128830..128951,
        130687..130819,133701..133808,136004..136114,
        140002..140114,153676..155831)
        /gene="RORB"
        /locus_tag="RP11-133M9.1-001"
        /standard_name="OTTHUMP0000021486"
        /note="match: proteins: P45446 Q8R1B8 Q92753 Q98934"
        /codon_start=1
        /product="RAR-related orphan receptor B"
        /protein_id="CAD13276.2"
        /db_xref="GI:55661490"
        /db_xref="GOA:Q92753"
        /db_xref="HGNC:10259"
        /db_xref="HSSP:1NQ7"
        /db_xref="InterPro:IPR000536"
        /db_xref="InterPro:IPR001628"
        /db_xref="InterPro:IPR001723"
        /db_xref="InterPro:IPR008946"
        /db_xref="UniProtKB/Swiss-Prot:Q92753"
        /translation="MRAQIEVLPCKICGDSXSGHYGVITCEGCKGFFRRSQNNASY
        SCPRQCLIDRTNRNCOHCRLOKIALGNSRDVAFGRMKQRDLSLAEVQKHQQ
        RLQROQOQSEALARYSSISNLSNLTSTGYANGHVIDLPKSEGVNDS
        IQSPDQSGDLMTKIQKQPIYDLTSVFNLTYSFNNQALPGATITTEIDRIQN
        I1KSHLTQVTEMLHQLAWOTHTYEIKAYQSKREALWQCAQI0ITHAIOVVFV
        AKRTIGFMELCONDIQLLSKSGCLEVLMRCFAFNLTNNLTVLFEGYKGMQMFKALG
        SDDLINAEAFPAKNCLSQLTEEBIALFSSAVLSIPDRAWLLIEPRKVQKLEQYFAL

```

QHVIOXNHLDDDELAKLIAKIPIITAVCNLHGEKLVQFKQSHPEIVNTLPFPLYKELF
 NPDCAVGTCK"
 154145
 /gene="RORB"
 /locus_tag="RP11-133M9.1-001"
 155384..155389
 /gene="RORB"
 /locus_tag="RP11-133M9.1-001"
 155406
 /gene="RORB"
 /locus_tag="RP11-133M9.1-001"
 157168
 /note="Clone right end: RP11-133M9"
 ORIGIN
 Query Match 13.3%; Score 40; DB 5; Length 157168;
 Best Local Similarity 52.4%; Pred. No. 0.057;
 Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 Qy 74 CTCCTCAATTTGTTACCTTAACCTGGAGTTTATCATGAAGAAATGTTTAATGCTTAATTT 133
 Db 30528 CTACTCCATTTGTTACCTTAACCTTTGAGATACCAAGATAATAATTTGAGCATATCTCTTATAGG 30469
 Qy 134 TCAATAAGAGTAACACAGAGTCTTCCACATCTGTTTAATAACCCGACAACTTCAAT 193
 Db 30468 TTAGATACGGAGTTTACATCTGACATATTAATGAGTGTACATAAGAAAGTAGACTGAAT 30409
 Qy 194 CACTATAGCTGTAGTAGAGTGCTTCTTCAAGGATCCAGAGTAACCA 241
 Db 30408 ACATACAGCATTTGATATAGAAGTTTCAAGGGTTCAAGGGTAGCAA 30361
 RESULT 34
 AC172335
 LOCUS
 DEFINITION
 AC172335.2 GI:85678238
 HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 Bos taurus clone CH240-265N2, WORKING DRAFT SEQUENCE, 10 unordered
 pieces.
 AC172335
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 160933)
 Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P.,
 Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
 Baraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
 Cadoree, J., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
 Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
 Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
 Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
 Anda, C., Delgado, O., Denson, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
 Donlin, J., McCauley, S., Dugan-Rocha, S., Espinosa, V., Eugene, C., Fa, M.,
 Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
 Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
 Gaskin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,
 Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haerberlen, K.,
 Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
 Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
 Hines, S., Hitchens, M., Hodgson, A., Hogue, J., Imo, K., Jackson, A., Jackson, L.,
 Howell, L., Hulyk, S., Hume, J., Imo, K., Johnson, R., Kalafus, K., Kelly, S.,
 Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kovar, C., Kowis, C., Lara, F.,
 Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Kowis, S.,
 Leal, S., Lee, K., Lee, S., Legall, F.I., Lemon, S., Lewis, L., Li, B.,
 Li, Y., Li, Z., Linell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyangde, D.,
 London, P., Lopez, J., Lorensuethu, L., Lozada, R., Luk, T., Madu, R.,
 Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
 McClelland, H., McPherson, J., Mercadado, C., Metzger, M.,
 Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
 Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,

JOURNAL Submitted (03-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Apr 5, 2004 this sequence version replaced gi:37992132.

COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhihong Bao and Sean Eddy, submitted) and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEYP-75B4 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES
source
1..181272
/location=Qualifiers
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clones="DKEYP-75B4"
/clone_lib="daniokeypilot"

ORIGIN
Query Match 13.3%; Score 40; DB 11; Length 181272;
Best Local Similarity 59.8%; Pred. No. 0.057;
Matches 67; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 23 AGTGACTACGATGAAGACTTGAATATCTTCATTGAAATAACAGTAACTCCCAAT 82
|||||
Db 43366 AGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 43425
|||||

Qy 83 TGTACCTCAACTAGGCGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
|||||

Db 43426 AATACCTAAGCAGAAACTTTATGATGATGATGATGATGATGATGATGAT 43477
|||||

RESULT 36
AC169312/c
LOCUS
DEFINITION Bos taurus clone CH240-256G15, *** SEQUENCING IN PROGRESS ***, 18
unordered pieces.
AC169312
AC169312.2 GI:85677635
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Bos taurus (cattle)
SOURCE
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 192411)
Muzny,D., Adams,C., Agbai II,O., Allen,C., Alebrooks,S., Archer,P., Arraondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R., Bereducci,A., Biswalo,K., Blyth,P., Bonham,H., Buhay,C., Burch,P., Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I., Chacko,J., Chahrour,M., Chavez,D., Chen,A., Chen,G., Chen,R., Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A., Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K., Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M., Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P., Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T., Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D., Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haerberlen,K., Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P., Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J., Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B., Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S., Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F., Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B., Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D., London,P., Lopez,J., Loresunewa,L., Lozado,R., Luk,T., Madu,R., Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E., McClelland,H., McPherson,J., Mercadao,C., Metzker,M., Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M., Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E., Nott,A., Nwokilemeh,O., Obregon,M., Ochi-Okorie,C., Odeh,E., Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B., Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T., Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J., Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S., Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W., Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y., Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R., Trejos,Z., Usmani,K., Varco,C., Verdusco,D., Villanana,D., Virk,D., Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J., Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R., Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L., Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.A.

Unpublished
2 (bases 1 to 192411)
Worley,K.C.
Direct Submission
Submitted (04-OCT-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 192411)

REFERENCE
Bovine Genome Sequencing Consortium
Direct Submission
Submitted (24-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 24, 2006 this sequence version replaced gi:76880524.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: FOKB
 Center clone name: CH240-256G15
 ----- Summary Statistics -----
 Assembly program: Atlas 3.0;
 Consensus quality: 185578 bases at least Q40
 Consensus quality: 186886 bases at least Q30
 Consensus quality: 188093 bases at least Q20
 Estimated insert size: 189246; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 16328: contig of 16328 bp in length
 * 16329 16378: gap of 50 bp
 * 16379 27546: contig of 11168 bp in length
 * 27547 27596: gap of 50 bp
 * 27597 36718: contig of 9122 bp in length
 * 36719 36768: gap of 50 bp
 * 36769 41738: contig of 4970 bp in length
 * 41739 41798: gap of 60 bp
 * 41799 45763: contig of 3965 bp in length
 * 45764 45813: gap of 50 bp
 * 45814 53973: contig of 8060 bp in length
 * 53974 53974: gap of unknown length
 * 53974 56584: contig of 2611 bp in length
 * 56585 57550: gap of 966 bp
 * 57551 64387: contig of 6837 bp in length
 * 64388 64437: gap of 50 bp
 * 64438 93575: contig of 29138 bp in length
 * 93576 93625: gap of 50 bp
 * 93626 108540: contig of 14915 bp in length
 * 108541 108590: gap of 50 bp
 * 108591 147592: contig of 39002 bp in length
 * 147593 147642: gap of 50 bp
 * 147643 162381: contig of 14739 bp in length
 * 162382 162431: gap of 50 bp
 * 162432 176477: contig of 14046 bp in length
 * 176478 176539: gap of 62 bp
 * 176540 177541: contig of 1002 bp in length
 * 177542 177641: gap of unknown length
 * 177642 185984: contig of 8343 bp in length
 * 185985 187209: gap of 1225 bp
 * 187210 189990: contig of 2781 bp in length
 * 189991 190091: gap of unknown length
 * 190091 191091: contig of 1001 bp in length
 * 191092 191191: gap of unknown length
 * 191192 192411: contig of 1220 bp in length.
 Location/Qualifiers
 1. 192411
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-256G15"
 16329. 16378
 /estimated_length=50
 27547. 27596
 /estimated_length=50
 36719. 36768
 /estimated_length=50
 41739. 41798

gap /estimated_length=60
 45764. 45813
 /estimated_length=50
 53874. 53973
 /estimated_length=unknown
 56585. 57550
 /estimated_length=966
 64388. 64437
 /estimated_length=50
 93576. 93625
 /estimated_length=50
 108541. 108590
 /estimated_length=50
 147593. 147642
 /estimated_length=50
 162382. 162431
 /estimated_length=50
 176478. 176539
 /estimated_length=62
 177542. 177641
 /estimated_length=unknown
 185985. 187209
 /estimated_length=1225
 189991. 190090
 /estimated_length=unknown
 191092. 191191
 /estimated_length=unknown
 ORIGIN
 Query Match 13.3%; Score 40; DB 12; Length 192411;
 Best Local Similarity 57.0%; Pred. No. 0.056;
 Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 6 TGACAGGCTAAATGCTAAGTACGATGACGATGAGGAGTTCGAAATATCTTCATTGGAATAA 65
 Db 118110 TTACAGGATAATAGGATAATCTGAGGGAAGAAACTTCCCAAGTCAGCCTCCCAAAATCA 118051
 QY 66 ACAGCTAACTCCCAAAATGTACTAACTAGGGAGTTCATCATGAAGAAATGTTTAAATG 125
 Db 118050 ATAGCAAGTGCCTACATACTAAACCTTAACAAGTGCTATTATGTTTATCATGATTCATT 117991
 QY 126 CTTAATTT 133
 Db 117990 TGAATAAT 117983
 RESULT 37
 ABI93599 Plasmodium falciparum rhoph1/clag8 mRNA linear INV 28-JUL-2005
 LOCUS
 DEFINITION
 cds.
 ACCESSION ABI93599 GI:71361251
 VERSION ABI93599.1
 KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)
 SOURCE Plasmodium falciparum
 ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1
 AUTHORS Kaneko, O., Yim Lim, B.Y.S., Iriko, H., Ling, I.T., Otsuki, H., Grainger, M., Tsuboi, T., Adams, J.H., Mattei, D., Holder, A.A. and Torii, M.
 TITLE Apical expression of three Rhoph1/Clag proteins as components of the Plasmodium falciparum Rhoph complex
 JOURNAL Mol. Biochem. Parasitol. 143 (1), 20-28 (2005)
 PUBMED 15953647
 REFERENCE 2 (bases 1 to 4250)
 AUTHORS Kaneko, O., Iriko, H., Tsuboi, T. and Torii, M.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2004) Osamu Kaneko, Ehime University School of Medicine, Department of Molecular Parasitology; Shitsukawa, Toon, Ehime, 791-0295, Japan (E-mail:okaneko@ehime-u.ac.jp, URU: <http://www.m.ehime-u.ac.jp/school/parasitology/eng/index.html>, Tel:81-89-960-5286, Fax:81-89-960-5287)

Assembly program: Phrap; version 0.990329
Consensus quality: 140365 bases at least Q40
Consensus quality: 143911 bases at least Q30
Consensus quality: 146396 bases at least Q20
Estimated insert size: 143156; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 33 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1137: contig of 1137 bp in length
* 1138 1237: gap of unknown length
* 1238 2283: contig of 1046 bp in length
* 2284 2383: gap of unknown length
* 2384 4115: contig of 1732 bp in length
* 4116 4215: gap of unknown length
* 4216 5435: contig of 1220 bp in length
* 5436 5535: gap of unknown length
* 5536 6746: contig of 1211 bp in length
* 6747 8104: gap of unknown length
* 8105 8204: contig of 1258 bp in length
* 8205 9949: contig of 1745 bp in length
* 9950 10049: gap of unknown length
* 10050 11148: contig of 1099 bp in length
* 11149 12657: contig of 1409 bp in length
* 12658 12757: gap of unknown length
* 12758 15248: gap of unknown length
* 15249 17176: contig of 1829 bp in length
* 17177 19265: contig of 1989 bp in length
* 19266 21679: gap of unknown length
* 21680 21779: gap of unknown length
* 21780 24338: contig of 2459 bp in length
* 24339 28792: contig of 4454 bp in length
* 28793 28892: gap of unknown length
* 28893 31953: contig of 3060 bp in length
* 31954 32052: gap of unknown length
* 32053 35524: contig of 3472 bp in length
* 35525 37842: gap of unknown length
* 37843 40865: contig of 2923 bp in length
* 40866 43954: gap of unknown length
* 43955 49162: contig of 2989 bp in length
* 49163 49262: gap of unknown length
* 49263 53030: contig of 5108 bp in length
* 53031 53131: gap of unknown length
* 53132 57858: gap of unknown length
* 57859 62862: contig of 5004 bp in length
* 62863 69621: gap of unknown length
* 69622 77621: gap of unknown length
* 77622 77662: contig of 8040 bp in length
* 77663 86146: gap of unknown length
* 86147 96279: contig of 8284 bp in length
* 96280 96379: gap of unknown length

96380 105100: contig of 8721 bp in length
* 105101 105200: gap of unknown length
* 105201 116139: contig of 10939 bp in length
* 116140 116239: gap of unknown length
* 116240 129563: contig of 13324 bp in length
* 129564 129663: gap of unknown length
* 129664 138778: contig of 9115 bp in length
* 138779 138879: gap of unknown length
* 138880 157011: contig of 18133 bp in length.
Location/Qualifiers
1. 157011
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-303C17"
1138. 1237
/estimated_length=unknown
2284. 2383
/estimated_length=unknown
4116. 4215
/estimated_length=unknown
5436. 5535
/estimated_length=unknown
6747. 6846
/estimated_length=unknown
8105. 8204
/estimated_length=unknown
9950. 10049
/estimated_length=unknown
11149. 11248
/estimated_length=unknown
12658. 12757
/estimated_length=unknown
15248. 15347
/estimated_length=unknown
17177. 17276
/estimated_length=unknown
19266. 19365
/estimated_length=unknown
21680. 21779
/estimated_length=unknown
Query Match 13.3%; Score 39.8; DB 12; Length 157011;
Best Local Similarity 46.3%; Pred. No. 0.066;
Matches 131; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
Qy 17 ATGCTAGTACTAGTACGATGAGACTTGAATATCTTCATTGAAATTAACAGCTAATCTC 76
Db 23417 ATGTGAAGGTTATTAAGTAAATAATTTAGGGTTTTTTTAAAGATTTATATAATATGATTTT 23476
Qy 77 CCAATTTGTACCTAACCTAGGGAGCTTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCA 136
Db 23477 TGGGAGAAGATTTTAAATAAATTTTATTTATATAAATAGAATGTTAGTTAGTTTAG 23536
Qy 137 AATAAGAAGTAACACAGAGTGTTCACATGCTGTTAAATTAACCCGACAACTTCAATCAC 196
Db 23537 AAAAAAGCGTAAATTTGTTATTAATAAATAAGTTAAGAAATAATTAATAAATAAATA 23596
Qy 197 TATAGCTGTAGTAGAGTGCATCTGCAAGATCCACAGTAACCAAGTATTTTGAATGC 256
Db 23597 AAGAATAAGTTTAAATGTTATTTTGAAGAAGTAAGTAAATTAATAAATTTTGAATTA 23656
Qy 257 AATGTTCAACCGACCATCACTAATTTATCTGCTGATTAGGAAGTA 299
Db 23657 AATAATAAATTTATGTTTAAAGTAATTAATTAATAAATA 23699
RESULT 39
AL929174/c
LOCUS AL929174 161610 bp DNA linear VRT 07-NOV-2003
DEFINITION Zebrafish DNA sequence from clone CH211-117C16 in linkage group 12,
complete sequence.
ACCESSION AL929174


```

TITLE Direct Submission
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 185657)
TITLE Worley,K.C.
JOURNAL Direct Submission
SUBMITTED (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 185657)
AUTHORS Rat Genome Sequencing Consortium.
JOURNAL Direct Submission
SUBMITTED (21-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 21, 2002 this sequence version replaced gi:22772818.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.
-----
Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GDPI
Center clone name: CH230-9P18
-----
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 169494 bases at least Q40
Consensus quality: 172042 bases at least Q30
Consensus quality: 173666 bases at least Q20
Estimated insert size: 184697; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 181626: contig of 181626 bp in length
* 181627 181726: gap of unknown length
* 181727 185657: contig of 3931 bp in length.
-----
FEATURES
source
Location/Qualifiers
1..185657
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9P18"
1..1734
/note="wgs_end_extension"
clone_end:T7"
1785..3006
/note="wgs_end_extension"
clone_end:T7"
6683..7563
/note="clone_boundary"
clone_end:T7"
site:EcoRI
end_sequence:BH344419"
76575..77257
/note="clone_boundary"
clone_end:Sp6
site:EcoRI
misc_feature
end_sequence:BH344422"
176639..177753
/note="wgs_end_extension"
clone_end:Sp6"
178823..180024
/note="wgs_end_extension"
clone_end:Sp6"
180462..181626
/note="wgs_end_extension"
clone_end:Sp6"
181627..181726
/estimated_length=unknown
ORIGIN
Query Match 13.3%; Score 39.8; DB 12; Length 185657;
Best Local Similarity 60.7%; Pred. No. 0.066;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 39 ACTTGAATATTCCTTCATTTGAATATAACAGCTAACTCCCAATTTGTACTTAAGGGG 98
Db 95656 ACTAGAAAGGTAAATTTTTTTTAAAAAATAACCTTACATCCACTGTGTTTCAGAAATGAGCAG 95597
QY 99 AGTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAATAAGAG 145
Db 95596 TGTCAATTTCTGACAAATAATACAAATTTGTAATTTTCTCAAAAGAG 95550
-----
RESULT 41
LOCUS BX005210 207949 bp DNA linear VRT 16-APR-2005
DEFINITION Zebrafish DNA sequence from clone DKEY-222F8 in linkage group 4,
complete sequence.
ACCESSION BX005210
VERSION BX005210
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS Pelan,S.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfzfsh-help@sanger.ac.uk Clone requests: Clonerequest@sanger.ac.uk
COMMENT On Apr 17, 2005 this sequence version replaced gi:41016365.
-----
Center: Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfzfsh-help@sanger.ac.uk
-----
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
DKEY-222F8 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml.

FEATURES

```

source
1. 207949
   /organism="Danio rerio"
   /mol_type="genomic DNA"
   /db_xref="taxon:7955"
   /clones="DKEY-222P8"
   /clone_lib="DanioKey"

```

ORIGIN

```

Query Match      13.3%; Score 39.8; DB 11; Length 207949;
Best local Similarity 50.8%; Pred. No. 0.066;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

34  TGAAGACTTGAATATCTTCATTGTGAATAAACAGCTAACTCCCAATTTGTACCTAACT 93
Dbb  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69897 TAAAACTAGAAATATGCTATCTACTACAAATAGAAATACATCTCACTTTAAGTTGATATC 69956

94  AGGGGAGTTTATCATGATGAAGAAATGTTTAAATGCTTTAATTTTCAATAAGAAGTAACAGA 153
Dbb  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69957 CCTGCAGATTAACTGAAATCATGACCTACTATTATTTCCCTTTTGCAGTTCACITGA 70016

154  GTGTGTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCATCATATAGCTGTAGTAGAGT 213
Dbb  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70017 ATAATTTGCGAGGCTAACAAATGAATTTTAAACACACAAAACCTCTCAATAGATGGGAAT 70076

214  GCATTCT 220
Dbb  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70077 GCAGACT 70083

```

```

RESULT 42
LOCUS      AC126085          208337 bp      DNA      linear      HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-14504, WORKING DRAFT SEQUENCE.
ACCESSION  AC126085
VERSION    AC126085.4  GI:25138258
KEYWORDS   HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE     Rattus norvegicus (Norway rat)

```

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 208337)

Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, I., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, F., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheewa, L., Loulaseg, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 208337)

Worley, K. C.

Direct Submission

Submitted (03-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 208337)

Rat Genome Sequencing Consortium

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:22856016.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZCO
Center clone name: CH230-14504
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 200812 bases at least Q40
Consensus quality: 202048 bases at least Q30
Estimated insert size: 202703 bases at least Q20
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 208337: contig of 208337 bp in length.
 Location/Qualifiers

FEATURES

source
 1. 208337
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-14504"

misc_feature

1. 1006
 /note="wgs_contig"

ORIGIN

Query Match 13.3%; Score 39.8; DB 12; Length 208337;
 Best Local Similarity 60.7%; Pred. No. 0.066; Mismatches 42; Indels 0; Gaps 0;
 Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 QY 39 ACTTGAATATCTTCATTGGAATTAACAGCTAACTCCAAATGTGTACCTAACTAGGGG 98
 DB 184783 ACTGAAAGGTAATTTTAAATAAATAACCTTACATCCACTGTGTTCAGATGAGCAG 184842
 QY 99 AGTTTATCATGAGAAATGTTTAATGCTTAATTTTCAATTAAGAG 145
 DB 184843 TGTCAATTTCTGAACAAATTAACAATGTGTAATTTCTCAAGAAG 184889

RESULT 43

AC157740 46494 bp DNA linear HTG 25-FEB-2005
 LOCUS
 Xenopus tropicalis clone ISB1-43N23, WORKING DRAFT SEQUENCE, 2
 DEFINITION
 unordered pieces.
 AC157740
 AC157740.1 GI:60279812
 VERSION
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS
 Xenopus tropicalis (Silurana tropicalis)
 SOURCE
 ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Unpublished
 JOURNAL
 Unpublished
 2 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Direct Submission
 TITLE
 Submitted (25-FEB-2005) Production Genomics Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
 94598-1698, USA

REFERENCE

1 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Unpublished
 JOURNAL
 Unpublished
 2 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Direct Submission
 TITLE
 Submitted (25-FEB-2005) Production Genomics Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
 94598-1698, USA

AUTHORS

Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

COMMENT

Project Information
 Center Project Name: 2789798
 Center clone name: ISB-43N23
 Summary Statistics
 Consensus quality: 45777 bases at least Q40
 Consensus quality: 46121 bases at least Q30
 Consensus quality: 46222 bases at least Q20
 Estimated insert size: 75000; agarose-fp estimation
 Estimated insert size: 46394; sum-of-contigs estimation
 Quality coverage: 8.24 in Q20 bases; agarose-fp estimation

Quality coverage: 13.32 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 17494: contig of 17494 bp in length
 * 17495 17594: gap of unknown length
 * 17595 46494: contig of 28900 bp in length.

FEATURES

Location/Qualifiers
 1. 46494
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:8364"
 /clone="ISB1-43N23"
 /clone_lib="ISB Xenopus tropicalis BAC library"
 17495..17594
 /estimated_length=unknown

gap

ORIGIN

Query Match 13.2%; Score 39.6; DB 12; Length 46494;
 Best Local Similarity 49.5%; Pred. No. 0.078; Mismatches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
 QY 79 AAATTTGTACCTAACCTAGGGAGTTTATCATGAAGAATGTTTAAATGCTTAATTTCAAA 138
 DB 13333 AAATTTGTATCATTCATGAATAATAATTAATGTAGATATTTCTAGGATTTAATGTAAT 13392
 QY 139 TAAGAAGTAACAGAGTGTTCACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
 DB 13393 GAACCTTTGCCAAAGTGTGCTCAATTTATTTAAAGAAAGAGAGAGACTGTAACTGTT 13452
 QY 199 TAGCTGTAGTAGTGCAATCTGCAAGATCCCGAGATACCAAGTATTTTGGAAATGCAA 258
 DB 13453 GCACCCCTTTATTGTGGAATAAAGATGCACAGAAAGTAAACATTAAACATGCAATCTGC 13512
 QY 259 TGTTGAACCGACCATCACTAATTTATCT 284
 DB 13513 CATCCACAGACCTAAATAAATGTAT 13538

RESULT 44

AC133155.2
 WPCOMMENT

Sequence split into 4 fragments LOCUS AC133155 Accession AC133155
 Fragment Name Begin End
 AC133155.0 1 110000
 AC133155.1 100001 210000
 AC133155.2 200001 310000
 AC133155.3 300001 398299

Continuation (3 of 4) of AC133155 from base 200001 (AC133155 Mus musculus clone RP24-418)

Query Match 13.2%; Score 39.6; DB 12; Length 110000;
 Best Local Similarity 52.4%; Pred. No. 0.077; Mismatches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 45 AATATTCTTCATTGGAATAAAGCTAACCTCCCAATTTGTACCTAAGCTAGGGAGTTTA 104
 DB 37295 AACTTTCTTTTATTGAAATAAAGGACACAAAGATGGGTGGCTAGGAAAGGGGTGGA 37354

QY 105 TCATGGAAGAATGTTTAAATGCTTAATTTCAATAAGAGTACCAAGAGTGTTCACACA 164
 DB 37355 TCTAGGAGGAGTCTGGGAAGGGGTAATGTGATAAAAAATGTTATGAGGAGGTTTCAAAA 37414

QY 165 TGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAG 210

DB 37415 AACTGATAATATGAGGAGGGAACAAATAAATAATGCAAGTTTATG 37460

RESULT 45

BX511308	123109 bp	DNA	linear	VRT 04-SEP-2004		
LOCUS	Zebrafish DNA sequence from clone DKXP-118E10 in linkage group 15, complete sequence.					
DEFINITION						
ACCESSION	BX511308					
VERSION	11	GI:51890185				
KEYWORDS	HTG.					
SOURCE	Danio rerio (zebrafish)					
ORGANISM	Danio rerio					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.					
REFERENCE	1 (bases 1 to 123109)					
AUTHORS	Gray,E.					
TITLE	Direct Submission					
JOURNAL	Submitted (04-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk					
	On Sep 4, 2004 this sequence version replaced gi:51469396.					
COMMENT	----- Genome Center					
	Center: Wellcome Trust Sanger Institute					
	Center code: SC					
	Web site: http://www.sanger.ac.uk					
	Contact: zfish-help@sanger.ac.uk					

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, ENMEL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep>

Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhiring Bao and Sean Eddy, submitted), and those beginning 'drv' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see <http://www.sanger.ac.uk/projects/D.erio/fishmask.shtml> DKEYP-118E10 is from a Zebrafish BAC library

```

FEATURES
  source
    location/Qualifiers
      1..123109
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /clone="DKEYP-118E10"
        /clone_lib="Daniokevoilot"

```

[illegible]

Db	65966	TGTCATGTTTATCATCTTGTTACCAATTTAATAAATGCCATGTCACAAAAAAGATTAAAA	66025
QY	153	AGTGTGGCAACATGCTGTTTAAATAACCCGACAAACTTC	190
Db	66026	AGGTTAGCTCAAAAGGTATTTAAGACCAAAATTAATCTC	66063
RESULT	46		
LOCUS	AC166491	154673 bp	DNA linear ROD 02-NOV-2005
DEFINITION	Mus musculus chromosome 15, clone RP24-380A15, complete sequence.		
ACCESSION	AC166491		
VERSION	AC166491.8	GI:78675901	
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 154673)		
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus chromosome 15, clone RP24-380A15		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 154673)		

REFERENCE	REFERENCE
AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
REFERENCE	REFERENCE
AUTHORS	AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 154673)

REFERENCE AUTHORS

Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckghalter, B., Camarata, J., Chang, J., Cheepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (02-NOV-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Nov 2, 2005 this sequence version replaced gi:77798120.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L33519

Center clone name: 380_A_15

FEATURES

source

Location/Qualifiers

1. 154673
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"
/clone="RP24-380A15"
/clone_lib="RPCI-24 Male Mouse BAC"
1381..1419
/rpt_family="AT_rich"
complement(3059..3399)
/rpt_family="ORR1D"
3644..3664
/rpt_family="AT_rich"
3676..3807
/rpt_family="B1_MM"
4024..4046
/rpt_family="B4"
4047..4175
/rpt_family="B1_MM"
4176..4199
/rpt_family="B4"
4507..4529
/rpt_family="T)n"
complement(4991..6108)
/rpt_family="L1_MM"
complement(6102..7005)
/rpt_family="L1"
complement(8198..8545)
/rpt_family="Lx2"
9547..9740
/rpt_family="B3"
9743..9796

repeat_region
/rpt_family="CA)n"
9827..9847
/rpt_family="AT_rich"
complement(10690..11695)
/rpt_family="Lx4"
11695..12379
/rpt_family="Lx4"
14806..14859
/rpt_family="CAAAA)n"
14913..15020
/rpt_family="B3A"
complement(15036..15159)
/rpt_family="Lx2"
15222..15337
/rpt_family="GA-rich"
15338..15554
/rpt_family="L1"
complement(15595..15717)
/rpt_family="RMER17A2"
15721..15812
/rpt_family="L1"
complement(16321..16436)
/rpt_family="MIR"
17416..17428
/rpt_family="B4A"
17429..17509
/rpt_family="B1_MM"
17525..17562
/rpt_family="CA)n"
17636..17712
/rpt_family="TA)n"
18802..18964
/rpt_family="TATAA)n"
19344..19586
/rpt_family="B3"
complement(19793..19865)
/rpt_family="MER99"
19936..19957
/rpt_family="AT_rich"
20138..20231
/rpt_family="CT-rich"
complement(20258..20551)
/rpt_family="Lx8"
complement(21400..21514)
/rpt_family="PB1"
21666..21701
/rpt_family="TG)n"
22616..22781
/rpt_family="RSINE1"
complement(22851..22961)
/rpt_family="Lx9"
23041..23107
/rpt_family="L1MC3"
23123..23234
/rpt_family="PB1D10"
23243..23346
/rpt_family="L1MC3"
23363..23614
/rpt_family="L1MC3"
23615..23652
/rpt_family="ATGTG)n"
23653..23769
/rpt_family="L1MC3"
25065..25125

Query Match 13.2%; Score 39.6; DB 6; Length 154673;

Best Local Similarity 52.4%; Pred. No. 0.077;

Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 45 AATATTTCTTCATTTGAATAAACAGTAACTCCCAATTTGTACCTACTAGGGGAGTTTA 104

Db 87286 AACTTTCTTTTATTTGAATAAAAAAGGACACAAAGATGGGTGGGTAGAAAAGGGGGTGA 87345

QY 45 AATATTCTTCATTGAAATAAACAGCTAACTCCAAATGTACTCTAAGGGAGTTTA 104
DB 160528 AACTTCTTTTATGAAATAAAAAAGGCACAAAGATGGGTGGCTAGGAAAGGGGTGGA 160587

ORIGIN

QY 45 AATATTCTTCATTGAAATAAACAGCTAACTCCAAATGTACTCTAAGGGAGTTTA 104
DB 160528 AACTTCTTTTATGAAATAAAAAAGGCACAAAGATGGGTGGCTAGGAAAGGGGTGGA 160587


```

Query Match      13.2%; Score 39.6; DB 6; Length 193188;
Best Local Similarity 50.5%; Pred. No. 0.077;
Matches 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 40 CTTGAATATTTCTTCATTGAAATTAACAGCTAACTCCCAATTTGTACCTAACCTAGGGGA 99
DB 75475 CTAGAAATGCTATGAATTTGTAATAGAGATTACACACAGACAGACCAATTAAGCAA 75416
QY 100 GTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAATTAAGAAGTAAACAGAGTGTG 159
DB 75415 TTTAAAGAGAGAAAATATGAGTTCTAGTTTTCAGAAAGAAAGAGATTGTAATTTATT 75356
QY 160 CAACATGCTGTTAAATTAACCGACAACTCACTCACTATAGCTGTAGTAGAGTGCAATTC 219
DB 75355 TAACTATTTTATATTAAAGAGCAAAATTTCAATTTATTGTAACGACGAGTAGCCTGTTGA 75296
QY 220 TGCAGGATC 229
DB 75295 CAGAAATATC 75286

RESULT 49
AC084288 200393 bp DNA linear ROD 31-DEC-2002
LOCUS Mus musculus Strain C57BL6/J Chromosome 2 BAC, RP23-108H5, complete
DEFINITION sequence.
AC084288
VERSION AC084288.9 GI:27436766
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
1 (bases 1 to 200393)
Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
High Throughput Mouse Sequencing
Unpublished
TITLE High Throughput Mouse Sequencing
JOURNAL
REFERENCE
AUTHORS Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
1 (bases 1 to 200393)
Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
Submitted (21-OCT-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
3 (bases 1 to 200393)
Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
Direct Submission
Submitted (29-JUL-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
4 (bases 1 to 200393)
Montgomery,K.T., Grills,G., Han,J., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,
Perera,A., Shim,C., Thomas,E., Brown,W.A. and Kucherlapati,R.
Direct Submission
Submitted (31-DEC-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
COMMENT On Dec 31, 2002 this sequence version replaced gi:22002114.
-----Genome Center:
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: http://www.hpcgg.org/Sequence/mouse.html
Contact: hpcg@medel.mgh.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this
clone unless otherwise noted. If there are overlapping clones, the
overlaps are noted in the beginning and end of the Features
listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550).
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Regions of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST
and cDNA sequences in Unigene. Genes demonstrate at least two exons
flanked by consensus splice sites that maintain sequence continuity
across the splice junctions. Sequences that are not identical
matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double
stranded sequence for all regions. All sequence is completed to a
standard of coverage with a minimum of 3 reads with no ambiguities.
If the sequence coverage for a region does not meet this standard,
it is indicated in the annotation as Low Coverage. Low coverage
linkages are verified by PCR product size verification or
verification of forward and reverse reads from clones which span
the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated average error rate is less than 1 per 10,000
bases using the Consed quality parameters. Regions that do not
meet this requirement are annotated as Low Quality.

-----Summary Statistics
Center project name: ADJ
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990329
Contig length: 200393
Fraction of Phrap value < 40: 0.001248
Error Rate in Consed: 0.01 per 10,000 bases
Number of N's in Consensus: 0

----- Distribution of Quality < 40 Bases:
# bases
1000
900
800
700
600
500
400
300
200
100
0
5 10 15 20 25 30 35 40
Phrap Value Range

FEATURES
source
Location/Qualifiers
1. 200393
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-108H5"
81..159
/rpt_family="(TG)n"
160..215
/standard_name="low coverage"
/notes="Low Coverage"
1156..1228
/rpt_family="MIR3"
8019..8356
/rpt_family="(TCTA)n"
complement(10415..10518)

repeat_region
misc_feature
repeat_region
repeat_region
repeat_region

```

```
repeat_region      /rpt_family="PB1D10"
10870. .10898
/rpt_family="AT_rich"
12332. .12455
/standard_name="Low coverage"
/note="Low coverage, Unsure Length, CT Repeat, size
confirmed by PCR"
repeat_region      12344. .12422
/rpt_family="TC)n"
15917. .15956
/rpt_family="TTTTG)n"
repeat_region      complement(16855. .16918)
/rpt_family="MER116"
18562. .18696
/rpt_family="MERVL"
repeat_region      complement(19493. .19642)
/rpt_family="MERSA"
19697. .20093
/rpt_family="WTC"
repeat_region      complement(20426. .20621)
/rpt_family="MIR"
22447. .22505
/rpt_family="CA)n"
repeat_region      complement(22557. .22943)
/rpt_family="WTD"
repeat_region      complement(23181. .23330)
/rpt_family="MIR"
25775. .25911
/rpt_family="MERSA"
repeat_region      27903. .28185
/rpt_family="B4A"
repeat_region      28219. .28299
/rpt_family="CA)n"
repeat_region      28418. .28612
/rpt_family="B3"
repeat_region      29438. .29535
/rpt_family="PB1D10"
repeat_region      29563. .29605
/rpt_family="Charlies"
repeat_region      complement(30231. .30457)
/rpt_family="URRIA"
repeat_region      complement(31394. .31501)
/rpt_family="MIR"
31665. .31689
/rpt_family="TA)n"
repeat_region      complement(32793. .32844)
/rpt_family="ID5"
repeat_region      complement(32845. .33071)
/rpt_family="URRIA"
repeat_region      complement(33072. .33128)
/rpt_family="ID5"
repeat_region      complement(33129. .33259)
/rpt_family="B1F"
repeat_region      complement(33260. .33387)
/rpt_family="MERSA"
33778. .33861
/standard_name="J05163"
repeat_region      34411. .34454
/rpt_family="TTTTG)n"
repeat_region      35355. .35390
/rpt_family="A-rich"
repeat_region      35447. .35573
/rpt_family="RMER17B"
36263. .36403
/standard_name="AW146109"
repeat_region      36629. .36654
/rpt_family="TTTTG)n"
repeat_region      complement(36710. .36777)
/rpt_family="ID_B1"
repeat_region      36783. .36834
/rpt_family="TC)n"
repeat_region      complement(36886. .37134)
/rpt_family="B4A"
```

```
repeat_region      complement(37135. .37195)
/rpt_family="ID_B1"
repeat_region      37440. .37643
/rpt_family="ID_B1"

Query Match      13.2%; Score 39.6; DB 6; Length 200393;
Best Local Similarity 50.5%; Pred. No. 0.077;
Matches 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy      40      CTTGAAATATTCTTCATTGAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGA 99
Db      10759    CTAGAAAGCTATGAATTTGCAATAGAGAAATTCACACAAGACACACAAATTAAGCA 10818
Qy      100      GTTATCATGGAAGAAATGTTTAAATGCTTAATTTTCAATTAAGAAGTAACACAGAGTGTG 159
Db      10819    TTTAAAGACAGAAAAAATATGAGTCTCTAGTTTTCGAAGAAAGAAAGATTGTAATTAAT 10878
Qy      160      CAACATGCTGTTAAATTAACCGACAACTTCAATCACTATAGCTGTAGTAGTGCATTC 219
Db      10879    TAACTATTTTATATTTTAAAGAGCAAAATTTCAATTTATTGTAACAGCAGTAGCTGTTGA 10938
Qy      220      TGCAAGGATC 229
Db      10939    CAGAAATATC 10948

RESULT 50
BX908768/c      229628 bp      DNA      linear      HTG 27-MAR-2004
LOCUS      BX908768
DEFINITION      Mus musculus chromosome 2 clone RP24-139J5, 5 unordered pieces.
ACCESSION      BX908768
VERSION      BX908768.3      GI:45772194
KEYWORDS      HTG; HTGS PHASE1; HTGS CANCELLED.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE      1      (bases 1 to 229628)
AUTHORS      Sims,S.
TITLE      Direct Submission
JOURNAL      Submitted (26-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 27, 2004 this sequence version replaced gi:41392784.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bn139J5
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 166190 bases at least Q40
Consensus quality: 166402 bases at least Q30
Consensus quality: 166549 bases at least Q20
Insert size: 229228; sum-of-contigs
Insert size: 172554; 6.7% error; agarose-fp
Quality coverage: 7.46x in Q20 bases; sum-of-contigs Quality
coverage: 10.06x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
```

* as soon as it is available and the accession number will
* be preserved.

1 154697: contig of 154697 bp in length
* 154698 154797: gap of 100 bp
* 154798 157100: contig of 2303 bp in length
* 157101 157200: gap of 100 bp
* 157201 211216: contig of 54016 bp in length
* 211217 211316: gap of 100 bp
* 211317 220309: contig of 8993 bp in length
* 220310 220409: gap of 100 bp
* 220410 229628: contig of 9219 bp in length.

FEATURES

source

1. .229628
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP24-139J5"
/clone_lib="RPC1-24"

misc_feature

1. .154697
/note="assembly_fragment:03328"
clone_end:SP6
vector_side:left

misc_feature

154798..157100
/note="assembly_fragment:00059"

misc_feature

157201..211216
/note="assembly_fragment:03367"
fragment_chain:1

misc_feature

211317..220309
/note="assembly_fragment:00391"
fragment_chain:1

misc_feature

220410..229628
/note="assembly_fragment:00388"
fragment_chain:1
clone_end:T7
vector_side:right

ORIGIN

Query Match 13.2%; Score 39.6; DB 12; Length 229628;
Best Local Similarity 50.5%; Pred. No. 0.077;
Matches 96, Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 40 CTTGAAATATCTTCATTGAAATPAAACAGCTAACTCCCAAAATGTACCTAACTAGGGGA 99
Db 80862 CTAGATAGCTATGCAATTGTGAATAGAGAATTCACACACAGACACAAATTAAGCAA 80803
Qy 100 GTTTATCATGAGAAATGTTTAAATGCTTAATTTTCAATTAAGAAAGTAAACAGAGTGTG 159
Db 80802 TTTAAAGAGAGAAAATAATAGTCTCTAGTTTGCAGAGAAAGAAATGTTAAATATT 80743
Qy 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTC 219
Db 80742 TAACATTTTATATTTAAGAGCAATTTTCATTTATGTAACAGCAGTAGCCTGTTTGA 80683
Qy 220 TGAAGGATC 229
Db 80682 CAGAAATATC 80673

Search completed: July 17, 2006, 22:34:45
Job time : 2391 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:41:41 ; Search time 386 Seconds
(without alignments)
5418.846 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1.9gcattgacaggctaaatgc.....atctgctgattaggaagtat 300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : N_Geneseq_8.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005as.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	299.6	100.0	110000	12	ADO79173_0	Ado79173 KLF12 gen
2	41.8	13.9	2000	11	ACL35363	ACL35363 Rice stre
3	39.4	13.1	60461	14	AED18414	Aed18414 Fibrotic
4	38.8	12.9	295096	11	ACN44068	Acn44068 Mouse gen
5	38.6	12.9	2224	5	ABV27665	Abv27665 Human pro
6	38.6	12.9	2224	5	ABV21841	Abv21841 Human pro
7	38.6	12.9	2224	5	ABV24590	Abv24590 Human pro
8	38.6	12.9	2224	5	ABV28008	Abv28008 Human pro
9	38.6	12.9	2224	5	ABV23184	Abv23184 Human pro
10	38.6	12.9	2224	5	ABV24513	Abv24513 Human pro
11	38.6	12.9	2224	5	ABV24601	Abv24601 Human pro
12	38.6	12.9	2224	5	ABV29023	Abv29023 Human pro
13	38.6	12.9	2224	5	ABV28014	Abv28014 Human pro
14	38.6	12.9	2224	5	ABV22169	Abv22169 Human pro
15	38.6	12.9	2224	5	ABV22175	Abv22175 Human pro
16	38.6	12.9	2224	5	ABV25057	Abv25057 Human pro
17	37.2	12.4	14920	6	ABN80147	Abn80147 Human che
18	37	12.3	6857	11	ADW53569	Adw53569 Human 98P

c 19	37	12.3	6857	11	ADW53577	Adw53577 Human 98P
c 20	37	12.3	6857	11	ADW53581	Adw53581 Human 98P
c 21	37	12.3	6857	11	ADW53696	Adw53696 Partial h
c 22	37	12.3	6857	11	ADW53561	Adw53561 Human 98P
c 23	37	12.3	6857	11	ADW53583	Adw53583 Human 98P
c 24	37	12.3	6857	11	ADW53579	Adw53579 Human 98P
c 25	37	12.3	6857	11	ADW53573	Adw53573 Human 98P
c 26	37	12.3	6857	11	ADW53575	Adw53575 Human 98P
c 27	37	12.3	6857	11	ADW53571	Adw53571 Human 98P
c 28	37	12.3	6857	11	ADW53698	Adw53698 Partial h
c 29	37	12.3	6857	11	ADW53585	Adw53585 Human 98P
c 30	37	12.3	6857	11	ADW53559	Adw53559 Human 98P
c 31	37	12.3	6857	11	ADW53563	Adw53563 Human 98P
c 32	37	12.3	6857	11	ADW53525	Adw53525 Human 98P
c 33	37	12.3	6857	11	ADW53565	Adw53565 Human 98P
c 34	37	12.3	6857	11	ADW53567	Adw53567 Human 98P
c 35	37	12.3	6857	13	ADR66608	Adr66608 Human pro
c 36	37	12.3	6857	13	ADR66266	Adr66266 Human pro
c 37	36.8	12.3	6857	14	AEC20351	Aec20351 cDNA enco
c 38	36.8	12.3	2272	4	AAH18058	Aah18058 Human cDN
c 39	36.8	12.3	2533	4	AAK84697	Aak84697 Human imm
c 40	36.8	12.3	2533	4	AAK84698	Aak84698 Human imm
c 41	36.6	12.2	82761	14	AED18524	Aed18524 Fibrotic
c 42	36.4	12.1	320	13	ACF86506	Acf86506 Human SIR
c 43	36.4	12.1	392	13	ACF88431	Acf88431 Human SIR
c 44	36.4	12.1	2374	15	AEF55932	Aef55932 Human gen
c 45	36.4	12.1	2374	15	AEF55696	Aef55696 Human gen
c 46	36.2	12.1	2000	11	ACL35363	ACL35363 Rice stre
c 47	36.2	12.1	9087	6	ABK31277	Abk31277 Signal tr
c 48	36.2	12.1	9087	6	ABL70238	AbL70238 Chemical
c 49	36.2	12.1	9087	6	AAK84697	Aak84697 Human imm
c 50	36.2	12.1	17738	6	ABL33538	AbL33538 Human imm
c 51	36.2	12.1	37992	13	ABD33572	Abd33572 Murine ca
c 52	36	12.0	6020	6	ABL32199	AbL32199 Human imm
c 53	36	12.0	6020	6	ABL49308	AbL49308 Human pol
c 54	35.8	11.9	2310	8	ABZ10134	AbZ10134 Haematopo
c 55	35.8	11.9	5310	10	ADE84160	Ade84160 Human lym
c 56	35.8	11.9	7588	4	AAK84697	Aak84697 Human imm
c 57	35.8	11.9	7588	6	ABL34012	AbL34012 Human imm
c 58	35.8	11.9	7588	6	ABK28409	Abk28409 DNA trans
c 59	35.6	11.9	2000	11	ACL35887	ACL35887 Rice stre
c 60	35.6	11.9	130263	6	ABK83573	AbK83573 Human cDN
c 61	35.6	11.9	277616	13	ABD32602	Abd32602 Human can
c 62	35.4	11.8	2381	5	AAK15800	Aak15800 Human six
c 63	35.4	11.8	3754	11	ADW53521	Adw53521 Human 98P
c 64	35.4	11.8	3754	11	ADW53684	Adw53684 Partial h
c 65	35.4	11.8	3754	11	ADW53686	Adw53686 Partial h
c 66	35.4	11.8	4329	5	AAK15801	Aak15801 Human ORF
c 67	35.4	11.8	23054	4	AAK84676	Aak84676 Human imm
c 68	35.4	11.8	37973	6	ABL34196	AbL34196 Human imm
c 69	35.4	11.8	99918	13	ABD33009	Abd33009 Human can
c 70	35.2	11.7	4220	6	ABN81677	Abn81677 Human CDK
c 71	35.2	11.7	4220	6	ABN81676	Abn81676 Human CDK
c 72	35.2	11.7	4233	6	AAK61454	Aak61454 Human gen
c 73	35.2	11.7	4233	8	ABZ10093	AbZ10093 Haematopo
c 74	35.2	11.7	4233	8	ABZ09989	Abz09989 Haematopo
c 75	35.2	11.7	4233	10	ADE84073	Ade84073 Human lym
c 76	35.2	11.7	110000	4	AAK95240_02	Continuation (3 of
c 77	35.2	11.7	110000	4	AAK96733_02	Continuation (3 of
c 78	35.2	11.7	110000	6	ABT00010_02	Continuation (3 of
c 79	35.2	11.7	110000	6	ABT01503_02	Continuation (3 of
c 80	35.2	11.7	110000	11	ADW70291_02	Continuation (3 of
c 81	35.2	11.7	110000	12	ADH77486_02	Continuation (3 of
c 82	35	11.7	476	5	ABV06229	Abv06229 Human pro
c 83	35	11.7	2000	11	ACL37108	ACL37108 Rice stre
c 84	35	11.7	6912	6	ABK28372	AbK28372 DNA trans
c 85	35	11.7	61020	4	AAK46788	Aak46788 Tumour su
c 86	35	11.7	110000	14	AAE04879_3	Continuation (4 of
c 87	34.8	11.6	1056	3	AAC59528	Aac59528 Human sec
c 88	34.8	11.6	2646	15	AEF55793	Aef55793 Human gen
c 89	34.8	11.6	18133	6	ABK40017	AbK40017 Human che
c 90	34.8	11.6	18133	6	ABL32940	AbL32940 Human imm
c 91	34.8	11.6	110000	5	AAI61373_4	Continuation (5 of

OY 78 CAAATTGTACCTAAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTTCAA 137
DB 27784 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAGATGAGTAAACAAATATGTAA 27725
OY 138 ATAAGACTACACAG 154
DB 27724 CTAATAAATAAGAGAG 27708

RESULT 4

ACN44068
ID ACN44068 standard; DNA; 295096 BP.

XX ACN44068;

DT 18-NOV-2004 (first entry)

DB Mouse genomic sequence MCG13636.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus;

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 331; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

XX Sequence 295096 BP; 75726 A; 52659 C; 64620 G; 85924 T; 0 U; 6167 Other;

XX Query Match 12.9%; Score 38.8; DB 11; Length 295096;

XX Best Local Similarity 48.9%; Pred. No. 4.3;

XX Matches 132; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

OY 1 GGCAATTCACAGGCTAAATGCTAAGTACTAGTACGAGACTTGAATATCTTCATTGA 60

DB 175844 GCCATTGCTGAAGACTCTGCTTGGGCTAGATGAAGAGATTCTATTAATTCCTTTGA 175903

OY 61 AATAACACAGCTAACTCCCAAAATTTACTACTAGGGAGTTTATCATGAAGAATGTTT 120

DB 175904 AAAGAAGAGCTCAAAACAGTCTCGTATTAATTTGTGTGTTTATCAAGTTCACTCTTA 175963

OY 121 AATGCTTAATTTTCAAAATAGAAAGTAACCAAGAGTGTTCACACATGCTGT-TAAATAACC 179

DB 175964 TGAGGAATGTTTTTAATAAAGAGAGTAAGCTGAGAAAGGAAAAATACAAAATATATGCT 176023
OY 180 CGACAAACTTCAATCACTATATAGCTGTAGTAGAGTGCATTTCTGCAAGGATCCACAGATTAAC 239
DB 176024 CAAGTATTAAAGGGGCATCAAGAAAGTAGAATGGAGCTGAATCCTATGTTCCAGGAGTTAA 176083
OY 240 CAGTATTTTGGAAATGCAATGTTTGAACCGA 269
DB 176084 CAAAATAAGGAGGTGTAATCTTGAGGCAA 176113

RESULT 5

ABV27665

ID ABV27665 standard; cDNA; 2224 BP.

XX AC ABV27665;

XX 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 27656.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5665; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

XX Query Match 12.9%; Score 38.6; DB 5; Length 2224;

XX Best Local Similarity 56.5%; Pred. No. 1.3;

XX Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

OY 41 TTGAAATATTCTTCATTGTAATAAACAGCTAACTCCAAATTTGTACTTAAGGGAG 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	2035	TTGAAATTATTTTAAAGCTTAAAAATCTGCTGGTTTGCAGAAATCTGATATATAACATTATATC	2094
Qy	101	TTTATCAAGAA-GAAATGTTTAAATGCTTAATTTTCAAATAAGAAAGTAACACAGAGTGTG	159
Db	2095	TATATTAATTAACGTGAAGTATTTTGGCTGATTAATGAATATATAAGTAAACACAGGGTTA	2154
Qy	160	CAACATGCTGTTAAATAACCCGCAAACTTCCAATCACTATA	200
Db	2155	CAGTTGTGCAAAAACAAAACAAAACATACCTAAATTACAATA	2195

RESULT 6	
ABV21841	ABV21841 standard; cDNA; 2224 BP.
XX	
XX	ABV21841;
XX	
DT	13-SEP-2002 (first entry)
XX	
DE	Human prostate expression marker cDNA 21832.
XX	
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW	pharmacogenomic marker; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200160860-A2.
XX	
PD	23-AUG-2001.
XX	
PF	20-FEB-2001; 2001WO-US005171.
XX	
PR	17-FEB-2000; 2000US-0183319P.
PR	16-MAR-2000; 2000US-0189862P.
PR	25-MAY-2000; 2000US-0207454P.
PR	09-JUN-2000; 2000US-0211314P.
PR	18-JUL-2000; 2000US-0219007P.
PR	13-DEC-2000; 2000US-0255281P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Schlegel R, Endege WO, Monahan JE;
XX	
DR	WPI; 2001-662795/76.
XX	
PT	Novel isolated nucleic acid molecule associated with cancerous state of
PT	prostate cells and correlating with presence of prostate cancer, useful
PT	for detecting presence of prostate cancer, stage of prostate cancer.
XX	
PS	Claim 1; Page 3700-3701; 117500pp; English.

Accession	Sequence	Position
Qy	101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAGAAAGTAAACCAAGAGTGTTG	159
Db	2095 TATATTATTAACTGTTAAGTATTTTGGCTGATTATGAAATATAAAGTAAACAACAGGGTTA	2154
Qy	160 CAACATGCTGTTAAATAACCCGCAAACTTCAATCACTATA	200
Db	2155 CAGTTGTGCAAAACAACAAACAAAACATACCTTAAATTTACAAATA	2195

RESULT 7
ABV24590
ID ABV24590 standard; cDNA; 2224 BP.
XX
XX AC ABV24590;
XX
XX
XX
DT 16-SEP-2002 (first entry)
XX
XX
XX Human prostate expression marker cDNA 24581.
XX
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WC200160860-A2.
FN
XX
XX
XX
XX 23-AUG-2001.
PD
XX
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX
XX 17-FEB-2000; 2000US-0183319P.
PR
XX 16-MAR-2000; 2000US-0189862P.
PR
XX 25-MAY-2000; 2000US-0207454P.
PR
XX 09-JUN-2000; 2000US-0211314P.
PR
XX 18-JUL-2000; 2000US-0219007P.
PR
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4665-4666; 11750pp; English.
PS

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACAGAGTGTG 159
DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTATGAATATAAAGTAAACAGGGTTA 2154
QY 160 CAACATGCTGTTTAAATACCCGACAAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAACAACATACCTTAAATTACAATA 2195

RESULT 8

ABV28008
ID ABV28008 standard; cDNA; 2224 BP.
XX
AC ABV28008;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 27999.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 5779; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
Best Local Similarity 56.5%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAAATATTTCTTCAATTTGAATAAACAAGCTAACTCCCAATTTGTAACCTAAGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCATTTGTAATAATATATC 2094
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACAGAGTGTG 159

DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTATGAATATAAAGTAAACAGGGTTA 2154
QY 160 CAACATGCTGTTTAAATACCCGACAAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAACAACATACCTTAAATTACAATA 2195

RESULT 9

ABV23184
ID ABV23184 standard; cDNA; 2224 BP.
XX
AC ABV23184;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23175.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 4163-4164; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
Best Local Similarity 56.5%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAAATATTTCTTCAATTTGAATAAACAAGCTAACTCCCAATTTGTAACCTAAGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCATTTGTAATAATATATC 2094
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACAGAGTGTG 159

Db	2095	TATATTATTAACTGCTAAAGTATTTTGGCTGATTAATGAATAATAAAGTAACAACAGGGTTA	2155
Qy	160	CAACATGCTGTTAAATAACCGCAAAACTTCAATCACTATA	200
Db	2155	CAGTTGTGCAAAACAAAACAAACATACCTAAATTTACAATA	2195
Db	160	CAACATGCTGTTAAATAACCGCAAAACTTCAATCACTATA	200
Db	2155	CAGTTGTGCAAAACAAAACAAACATACCTAAATTTACAATA	2195
RESULT 10			
ABV24513			
ID	ABV24513	standard; cDNA; 2224 BP.	
AC	ABV24513;		
AC	ABV24513;		
DT	16-SEP-2002	(first entry)	
XX	Human prostate expression marker cDNA 24504.		
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
KW	pharmacogenomic marker; gene; ss.		
KW	Homo sapiens.		
XX	WO200160860-A2.		
XX	23-AUG-2001.		
XX	20-FEB-2001; 2001WO-US005171.		
XX	17-FEB-2000; 2000US-0183319P.		
PR	16-MAR-2000; 2000US-0189862P.		
PR	25-MAY-2000; 2000US-0207454P.		
PR	09-JUN-2000; 2000US-0211314P.		
PR	18-JUL-2000; 2000US-0219007P.		
PR	13-DEC-2000; 2000US-0255281P.		
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
PA	Schlegel R, Endege WO, Monahan JE;		
XX	WPI; 2001-662795/76.		
XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.		
XX	Claim 1; Page 4637-4638; 11750pp; English.		
XX	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker		
XX	Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;		
XX	Query Match 12.9%; Score 38.6; DB 5; Length 2224;		
XX	Best Local Similarity 56.5%; Pred. No. 1.3;		
XX	Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;		
Qy	41	TTGAAATATCTTCTTATTGAAATAAACAGCTAACTCCCAATTCCTACTACTAGGGAG	100
Db	2035	TTGAAATATTTTAAAGCTTAAATTTCTGCTGGTTGCAAAATTTGTAATATCAATTATC	2094
Qy	101	TTTATCATGAA-GAAATGTTTTAAATGCTTAAATTTTCAATTAAGAGTAACACAGAGTTG	159
Db	2095	TATATTATTACTGTAAGTATTTTGGCTGATTTGAATATAAAGTAACAACAGGGTTA	2154

[illegible]

	Query Match	12.9%	Score 38.6;	DB 5;	Length 2224;
	Best Local Similarity	56.5%;	Prod. No. 1.3;		
	Matches 91;	Conservative 0;	Mismatches 69;	Indels 1;	Gaps 1;
QY	41	TTGAAATATTCTTCATTTGAAATAAACAAGCTAACTCCCAAATTGTACTTAAGTGGGAG	100		
DB	2035	TTGAATTAATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAATTGTATATACATTATATC	2094		
QY	101	TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAATAAGAAAGTAAACACAGAGTGTTG	159		
DB	2095	TATATATTAACTCTGAAGTATTTTTCCTGATATGAATATGAATATAAGTAAACAACAGGGTTA	2154		

Qy 160 CACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 ||
 Ds 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
 ||

RESULT 12
 ABV29023
 ID ABV29023 standard; cDNA; 2224 BP.
 XX
 AC ABV29023;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 29014.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PE 20-FEB-2001; 2001WO-US005171.
 XX
 PF 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
 Claim 1; Page 6143; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
 Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTCTTCAATTTGAATTAACAGCTAACTCCCAATTTGCTAACTAATCACTAGGGAG 100
 |||||
 Ds 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTGCAAAATTTGATATAATCACTATATC 2094
 |||||

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
 |||||
 Ds 2095 TATATTAATTAAGTAACTGTAATTTTGGCTGATTAATGAATTAAGTAACAGAGGTTA 2154
 |||||

Qy 160 CACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 ||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
 ||

RESULT 13
 ABV28014
 ID ABV28014 standard; cDNA; 2224 BP.
 XX
 AC ABV28014;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 28005.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PE 20-FEB-2001; 2001WO-US005171.
 XX
 PF 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
 Claim 1; Page 5782; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
 Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTCTTCAATTTGAATTAACAGCTAACTCCCAATTTGCTAACTAATCACTAGGGAG 100
 |||||
 Ds 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTGCAAAATTTGATATAATCACTATATC 2094
 |||||

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
 |||||
 Ds 2095 TATATTAATTAAGTAACTGTAATTTTGGCTGATTAATGAATTAAGTAACAGAGGTTA 2154
 |||||

Qy 160 CACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 ||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 14

ABV22169
ID ABV22169 standard; cDNA; 2224 BP.

XX
AC ABV22169;

XX
DT 13-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 22160.

XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.

XX
PN W0200160860-A2.

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001WO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

XX
PR 16-MAR-2000; 2000US-0189862P.

XX
PR 25-MAY-2000; 2000US-0207454P.

XX
PR 09-JUN-2000; 2000US-0211314P.

XX
PR 18-JUL-2000; 2000US-0219007P.

XX
PR 13-DEC-2000; 2000US-0255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
PI WPI; 2001-662795/76.

XX
DR Novel isolated nucleic acid molecule associated with cancerous state of

XX
PT prostate cells and correlating with presence of prostate cancer, useful

XX
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS Claim 1; Page 3811; 11750pp; English.

XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX
CC specification or its complement. (I) is useful for: (a) assessing whether

XX
CC a patient is afflicted with prostate cancer; (b) monitoring the

XX
CC progression of prostate cancer in a patient; (c) assessing the efficacy

XX
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX
CC determining whether prostate cancer has metastasized in a patient; (h)

XX
CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;

Best Local Similarity 56.5%; Pred. No. 1.3;

Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTAGGGGAG 100

Db 2035 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTATATATC 2094

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159

Db 2095 TATATTTAACTGTAAGTATTTTGGCTGATTAAGTAAGTAACAGAGGTTA 2154

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 15

ABV22175
ID ABV22175 standard; cDNA; 2224 BP.

XX
AC ABV22175;

XX
DT 13-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 22166.

XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.

XX
PN W0200160860-A2.

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001WO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

XX
PR 16-MAR-2000; 2000US-0189862P.

XX
PR 25-MAY-2000; 2000US-0207454P.

XX
PR 09-JUN-2000; 2000US-0211314P.

XX
PR 18-JUL-2000; 2000US-0219007P.

XX
PR 13-DEC-2000; 2000US-0255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
PI WPI; 2001-662795/76.

XX
DR Novel isolated nucleic acid molecule associated with cancerous state of

XX
PT prostate cells and correlating with presence of prostate cancer, useful

XX
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS Claim 1; Page 3812-3813; 11750pp; English.

XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX
CC specification or its complement. (I) is useful for: (a) assessing whether

XX
CC a patient is afflicted with prostate cancer; (b) monitoring the

XX
CC progression of prostate cancer in a patient; (c) assessing the efficacy

XX
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX
CC determining whether prostate cancer has metastasized in a patient; (h)

XX
CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;

Best Local Similarity 56.5%; Pred. No. 1.3;

Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTAGGGGAG 100

Db 2035 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTATATATC 2094

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159

Db 2095 TATATTTAACTGTAAGTATTTTGGCTGATTAAGTAAGTAACAGAGGTTA 2154

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 16
ABV25057
ID ABV25057 standard; cDNA; 2224 BP.
XX AC
XX ABV25057;
XX DT
XX 16-SEP-2002 (first entry)
XX DE
XX Human prostate expression marker cDNA 25048.
XX KW
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX KW
XX Homo sapiens.
XX OS
XX WO200160860-A2.
XX PN
XX 23-AUG-2001.
XX PD
XX 20-FEB-2001; 2001WO-US005171.
XX PF
XX 17-FEB-2000; 2000US-0183319P.
XX PR
XX 16-MAR-2000; 2000US-0189862P.
XX PR
XX 25-MAY-2000; 2000US-0207454P.
XX PR
XX 09-JUN-2000; 2000US-0211314P.
XX PR
XX 18-JUL-2000; 2000US-0219007P.
XX PR
XX 13-DEC-2000; 2000US-0255281P.
XX PA
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI
XX Schlegel R, Endege WO, Monahan JE;
XX DR
XX WPI; 2001-662795/76.
XX PT
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS
XX Claim 1; Page 4864; 11750pp; English.
XX CC
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (II) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX CC
XX Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
Query Match 12.9%; Score 38.6; DB 5; Length 2224;
Best Local Similarity 56.5%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTGTAATAAACAAGCTAACTCCCAATTTGCTACTAGGGAG 100
DB 2035 TTGAATATTCTTAAAGCTTAAATTTCTGCTGTTTGAATTTGATATACTATTATC 2094
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAATTAAGAAGTAACAGAGTGTG 159
DB 2095 TATATTTACTGTAAGTATTTTTCCTGATTATGAATATAAAGTAACAACAGGGTTA 2154
QY 160 CAACATGCTGTTAAATACCCGACAACTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAAACAAACATACCTTAAATTAACAATA 2195

RESULT 17
ABN80147/c
ID ABN80147 standard; DNA; 14920 BP.
XX AC
XX ABN80147;
XX DT
XX 15-JUL-2002 (first entry)
XX DE
XX Human chemically modified disease associated gene SEQ ID NO 164.
XX KW
XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
XX KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
XX KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
XX KW antidiabetic; cytostatic; anticonvulsant; ds.
XX KW
XX Homo sapiens.
XX OS
XX Synthetic.
XX PN
XX WO200200927-A2.
XX PD
XX 03-JAN-2002.
XX PF
XX 02-JUL-2001; 2001WO-EP007536.
XX PR
XX 30-JUN-2000; 2000DE-01032529.
XX PR
XX 01-SEP-2000; 2000DE-01043826.
XX PR
XX (EPIG-) EPIGENOMICS AG.
XX PA
XX Olek A, Piepenbrock C, Berlin K;
XX PI
XX WPI; 2002-130908/17.
XX DR
XX Novel nucleic acid useful for diagnosis and therapy of diseases
XX PT associated with development genes such as diabetes, comprises a sequence
XX PT of a segment of chemically pretreated DNA of genes associated with
XX PT development.
XX PS
XX Claim 1; SEQ ID NO 164; 27pp; English.
XX CC
XX The invention relates to a nucleic acid (I) comprising a sequence at
XX CC least 18 bases in length of a segment of chemically pretreated DNA (II)
XX CC of genes associated with development selected from 87 genes listed in the
XX CC specification such as ACCFN, ADFN, or AFD1 and comprising one of 350
XX CC sequences (ABN79984-ABN80333) or their complements. The invention is
XX CC useful for the diagnosis or therapy of diseases associated with
XX CC development genes, in particular disease related to homeobox containing
XX CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
XX CC associated with congenital heart disease, epilepsy, diseases related to
XX CC histone deacetylation, Currarino syndrome, diseases related with the
XX CC development of the brain and limb girdle muscular dystrophy and dwarfism.
XX CC Oligomers specific to each of the genes are useful for detecting the
XX CC methylation state of all CpG dinucleotides within the 350 sequences or
XX CC (II) and their complementary sequences, as primer oligonucleotides for
XX CC the amplification of the 350 sequences, (II) and/or their complements and
XX CC as oligomer probes for detecting the cytosine methylation state and/or
XX CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
XX CC patent did not form part of the printed specification but is based on
XX CC sequence information supplied to Derwent by the European Patent Office
XX CC
XX Sequence 14920 BP; 4132 A; 171 C; 3248 G; 7369 T; 0 U; 0 Other;
Query Match 12.4%; Score 37.2; DB 6; Length 14920;
Best Local Similarity 49.0%; Pred. No. 5.5;
Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 40 CTTGAATATTCTTCATTGCAATAAACAAGCTAACTCCCAATTTGCTACTAGGGGA 99
DB 12920 CATAAATAACAATAATTTTATATAAAACATATAAACTAACTAATATAACGAACGT 12861
QY 100 GTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATAAGAAGTAACAGAGTGTG 159
DB 12860 ATATACGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 12801

QY 160 CAACATGCTGTTAAATAAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATC 219
DB 12800 CAATATTTTCTAAAAATAAATAATCTCTTAACCAAAAAATTTCACTAACTCAATACATCA 12741
QY 220 TGCAGAGTCCAGAGTACCA 241
DB 12740 CCCAATAATATCTCAAAACCA 12719

RESULT 18
ADW53569/C
ID ADW53569 standard; cDNA; 6857 BP.
XX
AC ADW53569;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human 98P4B6 cDNA variant 30 SeqID 60.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosstatic; gene; chromosome 7q21.
XX
OS Homo sapiens.
XX
PN WO2003087306-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PR 06-JUN-2002; 2002US-00165044.
XX
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
WPI; 2003-903158/82.
XX
P-PSDB; ADW53570.

A composition comprising 98P4B6 proteins and nucleic acid molecules for detecting, preventing, prognosing and/or treating cancers that express 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
PS Example 2; SEQ ID NO 60; 616pp; English.
XX
CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2019 A; 1318 C; 1210 G; 2310 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGGAATATCTTCATTTGAAATAAACAAGCTAACTCCCAATTTGTACTTAAGGGAG 100

DB 2234 TTGAATATTTAAAAAGCTTAAAAATTCGTGCTTTCGCAATTTGTTTTTAACATTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTTAAATTTTCAAAATAAGTAACACAGAGTGTG 159
DB 2174 TATATTTTAACTGTAAAGTATTTTTCCTGATTTATGAATATAAAGTAACAACAGGTTA 2115
QY 160 CAACATGCTGTTAAATAAACCCGACAACTTCAATCACTATA 200
DB 2114 CAGTTGCAAAACAAACAAACATACCTTAATTTACAATA 2074

RESULT 19
ADW53577/C
ID ADW53577 standard; cDNA; 6857 BP.
XX
AC ADW53577;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human 98P4B6 cDNA variant 34 SeqID 68.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosstatic; gene; chromosome 7q21.
XX
OS Homo sapiens.
XX
PN WO2003087306-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PR 06-JUN-2002; 2002US-00165044.
XX
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
WPI; 2003-903158/82.
XX
P-PSDB; ADW53578.

A composition comprising 98P4B6 proteins and nucleic acid molecules for detecting, preventing, prognosing and/or treating cancers that express 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
PS Example 2; SEQ ID NO 68; 616pp; English.
XX
CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2018 A; 1317 C; 1210 G; 2312 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 41 TTGAATATCTTCAATTTGAATAAAGCAGTAACCTCCAAATTTGACCTAAGTAGGGGAG 100
DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCCTGTTTGAATTTGTTTAAACATTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAGAGTAACAGAGTGTG 159
DB 2174 TATATTTAATCTGTAAGTATTTTTCCTGCTGATTTGAATATATAAGTAACACAGGGTTA 2115
QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAAACATCACTTAAATTACAATA 2074

RESULT 20
ADW53581/c
ID ADW53581 standard; cDNA; 6857 BP.
XX
AC ADW53581;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human 98P4B6 cDNA variant 36 SeqID 72.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosolic; gene; chromosome 7q21.
XX
OS Homo sapiens.
XX
PN WO2003087306:A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
PR 06-JUN-2002; 2002US-00165044.
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
PI WPI; 2003-903158/82.
XX
DR P-PSDB; ADW53582.
XX
PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
PS Example 2; SEQ ID NO 72; 616pp; English.
XX
CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
CC 2, six transmembrane epithelial antigen of prostate-2) located on
CC chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult
CC normal tissues and its aberrant over-expression in various cancers
CC including breast, colon, lung, kidney and prostate. The present invention
CC describes compositions and methods useful for detecting, preventing,
CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2018 A; 1317 C; 1211 G; 2311 T; 0 U; 0 Other;

Query Match

12.3%; Score 37; DB 11; Length 6857;

Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATCTTCAATTTGAATAAAGCAGTAACCTCCAAATTTGACCTAAGTAGGGGAG 100
DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCCTGTTTGAATTTGTTTAAACATTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAGAGTAACAGAGTGTG 159
DB 2174 TATATTTAATCTGTAAGTATTTTTCCTGCTGATTTGAATATATAAGTAACACAGGGTTA 2115
QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAAACATCACTTAAATTACAATA 2074

RESULT 21
ADW53696/c
ID ADW53696 standard; cDNA; 6857 BP.
XX
AC ADW53696;
XX
DT 24-MAR-2005 (first entry)
XX
DE Partial human 98P4B6 cDNA variant 8 SeqID 187.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosolic.
XX
OS Homo sapiens.
XX
PN WO2003087306-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
PR 06-JUN-2002; 2002US-00165044.
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
PI WPI; 2003-903158/82.
XX
DR P-PSDB; ADW53699; ADW53701.
XX
PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
PS Example 51; SEQ ID NO 187; 616pp; English.
XX
CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
CC 2, six transmembrane epithelial antigen of prostate-2) located on
CC chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult
CC normal tissues and its aberrant over-expression in various cancers
CC including breast, colon, lung, kidney and prostate. The present invention
CC describes compositions and methods useful for detecting, preventing,
CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a partial human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
41 TTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTCTAACTAAGGGAG 100
2234 TTGAATATTTTAAAGCTTAAATTCCTGCTGGTTTGCATAATGTTTTTAACATATATC 2175
101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAAATAAGAAAGTAACAGAGTGTG 159
2174 TATATTATTAACCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2115
160 CAACATGCTGTTAAATAACCGCAAAACTTCAATCACTATA 200
2114 CAGTTGTGCAAAACAAACACATACCTAAATTAACAATA 2074

RESULT 22

ADW53561/c
ID ADW53561 standard; cDNA; 6857 BP.

XX AC ADW53561;

XX AC ADW53561 (first entry)

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 26 SeqID 52.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;

XX KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;

XX KW cytosstatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX PN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;

XX DR WPI; 2003-903158/82.

XX DR P-PSDB; ADW53562.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for

XX PT detecting, preventing, prognosing and/or treating cancers that express

XX PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 2; SEQ ID NO 52; 616pp; English.

XX CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-

XX CC 2, six transmembrane epithelial antigen of prostate-2) located on

XX CC chromosome 7q21 and the encoded protein and variants derived thereof.

XX CC Specifically, it refers to the expression pattern of this gene in adult

XX CC normal tissues and its aberrant over-expression in various cancers

XX CC including breast, colon, lung, kidney and prostate. The present invention

XX CC describes compositions and methods useful for detecting, preventing,

XX CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the

XX CC polynucleotide can be used for gene therapy purposes, for monitoring

XX CC genetic abnormalities and for generating transgenic or 'knockout'

XX CC animals, which can be useful for the development and screening of

XX CC therapeutically useful reagents. The encoded proteins may also be used in

XX CC generating antibodies and vaccines, as well as in identifying ligands and

XX CC cellular constituents that bind to 98P4B6 gene products. Accordingly,

XX CC these compositions exhibit cytostatic activities. This polynucleotide

XX CC sequence is a human 98P4B6 cDNA variant of the invention.

XX SQ Sequence 6857 BP; 2018 A; 1318 C; 1209 G; 2312 T; 0 U; 0 Other;
Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
41 TTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTCTAACTAAGGGAG 100
2234 TTGAATATTTTAAAGCTTAAATTCCTGCTGGTTTGCATAATGTTTTTAACATATATC 2175
101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAAATAAGAAAGTAACAGAGTGTG 159
2174 TATATTATTAACCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2115
160 CAACATGCTGTTAAATAACCGCAAAACTTCAATCACTATA 200
2114 CAGTTGTGCAAAACAAACACATACCTAAATTAACAATA 2074

RESULT 23

ADW53583/c

ID ADW53583 standard; cDNA; 6857 BP.

XX AC ADW53583;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 37 SeqID 74.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;

XX KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;

XX KW cytosstatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX PN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;

XX DR WPI; 2003-903158/82.

XX DR P-PSDB; ADW53584.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for

XX PT detecting, preventing, prognosing and/or treating cancers that express

XX PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 2; SEQ ID NO 74; 616pp; English.

XX CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-

XX CC 2, six transmembrane epithelial antigen of prostate-2) located on

XX CC chromosome 7q21 and the encoded protein and variants derived thereof.

XX CC Specifically, it refers to the expression pattern of this gene in adult

XX CC normal tissues and its aberrant over-expression in various cancers

XX CC including breast, colon, lung, kidney and prostate. The present invention

XX CC describes compositions and methods useful for detecting, preventing,

XX CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the

XX CC polynucleotide can be used for gene therapy purposes, for monitoring

XX CC genetic abnormalities and for generating transgenic or 'knockout'

XX CC animals, which can be useful for the development and screening of

XX CC therapeutically useful reagents. The encoded proteins may also be used in

XX CC generating antibodies and vaccines, as well as in identifying ligands and

XX CC cellular constituents that bind to 98P4B6 gene products. Accordingly,

XX CC these compositions exhibit cytostatic activities. This polynucleotide

XX CC sequence is a human 98P4B6 cDNA variant of the invention.

CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.

XX SQ Sequence 6857 BP; 2017 A; 1319 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTTGAAATTAACAGCTAACTCCCAATTTGTACTTAACCTAGGGAG 100
DB 2234 TTGAATATTCTTCATTTGAAATTAACAGCTAACTCCCAATTTGTACTTAACCTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACAACAGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATTAACCGACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACATACCTAAATTACAATA 2074

RESULT 26

ADW53575/c
ID ADW53575 standard; cDNA; 6857 BP.

XX AC ADW53575;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 33 SeqID 66.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
XX prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
XX cytostatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX EN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX P-PSDB; ADW53576.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 2; SEQ ID NO 66; 616pp; English.

XX PS This invention relates to a novel human gene 98P4B6 (also known as STEAP-
XX 2, six transmembrane epithelial antigen of prostate-2) located on
XX chromosome 7q21 and the encoded protein and variants derived thereof.
XX Specifically, it refers to the expression pattern of this gene in adult
XX normal tissues and its aberrant over-expression in various cancers
XX including breast, colon, lung, kidney and prostate. The present invention
XX describes compositions and methods useful for detecting, preventing,
XX prognosing and/or treating cancers that express 98P4B6. Furthermore, the

CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.

XX SQ Sequence 6857 BP; 2017 A; 1318 C; 1211 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTTGAAATTAACAGCTAACTCCCAATTTGTACTTAACCTAGGGAG 100
DB 2234 TTGAATATTCTTCATTTGAAATTAACAGCTAACTCCCAATTTGTACTTAACCTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACAACAGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATTAACCGACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACATACCTAAATTACAATA 2074

RESULT 27

ADW53571/c
ID ADW53571 standard; cDNA; 6857 BP.

XX AC ADW53571;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 31 SeqID 62.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
XX prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
XX cytostatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX EN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX P-PSDB; ADW53572.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 2; SEQ ID NO 62; 616pp; English.

XX PS This invention relates to a novel human gene 98P4B6 (also known as STEAP-
XX 2, six transmembrane epithelial antigen of prostate-2) located on
XX chromosome 7q21 and the encoded protein and variants derived thereof.
XX Specifically, it refers to the expression pattern of this gene in adult
XX normal tissues and its aberrant over-expression in various cancers
XX including breast, colon, lung, kidney and prostate. The present invention

describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2017 A; 1318 C; 1211 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
41 TTGAAATATCTTCATTTGAATAAAGCTTAAATCTCCCAATTTGCTAACTAGGAG 100
2234 TTGAATATTTTAAAGCTTAAATCTCCCAATTTGCTAACTAGGAG 2175
101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAAGTAACAGAGTGTG 159
2174 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATATAAGTAACAGAGGTTA 2115
160 CAACATGCTGTTAATAACCGACAACTTCAATCACTATA 200
2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 28
ADW53698/c
ID ADW53698 standard; cDNA; 6857 BP.

ADW53698;
24-MAR-2005 (first entry)
Partial human 98P4B6 cDNA variant 8 SeqID 189.
ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
cytostatic.
Homo sapiens.
W02003087306-A2.
23-OCT-2003.
04-APR-2003; 2003WO-US010462.
05-APR-2002; 2002US-0370387P.
06-JUN-2002; 2002US-00165044.
20-DEC-2002; 2002US-0435480P.
(AGEN-) AGENYSYS INC.
Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
WPI; 2003-903158/82.
P-PSDB; ADW53698, ADW53701.

A composition comprising 98P4B6 proteins and nucleic acid molecules for detecting, preventing, prognosing and/or treating cancers that express 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
Example 51; SEQ ID NO 189; 616pp; English.
This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof. Specifically, it refers to the expression pattern of this gene in adult

normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a partial human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
41 TTGAAATATCTTCATTTGAATAAAGCTTAAATCTCCCAATTTGCTAACTAGGAG 100
2234 TTGAATATTTTAAAGCTTAAATCTCCCAATTTGCTAACTAGGAG 2175
101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAAGTAACAGAGTGTG 159
2174 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATATAAGTAACAGAGGTTA 2115
160 CAACATGCTGTTAATAACCGACAACTTCAATCACTATA 200
2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 29
ADW53585/c
ID ADW53585 standard; cDNA; 6857 BP.

ADW53585;
24-MAR-2005 (first entry)
Human 98P4B6 cDNA variant 38 SeqID 76.
ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
cytostatic; gene; chromosome 7q21.
Homo sapiens.
W02003087306-A2.
23-OCT-2003.
04-APR-2003; 2003WO-US010462.
05-APR-2002; 2002US-0370387P.
06-JUN-2002; 2002US-00165044.
20-DEC-2002; 2002US-0435480P.
(AGEN-) AGENYSYS INC.
Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
WPI; 2003-903158/82.
P-PSDB; ADW53586.
A composition comprising 98P4B6 proteins and nucleic acid molecules for detecting, preventing, prognosing and/or treating cancers that express 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
Example 2; SEQ ID NO 76; 616pp; English.
This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on

CC chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult
CC normal tissues and its aberrant over-expression in various cancers
CC including breast, colon, lung, kidney and prostate. The present invention
CC describes compositions and methods useful for detecting, preventing,
CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2019 A; 1317 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTGAAATTAACAGCTAACTCCCAAATGTACTTAACCTAGGGGAG 100
DB 2234 TTGAATATTCTTCATTGAAATTAACAGCTAACTCCCAAATGTCTGCTGTTTGCATAATGTTTAAACATTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACCTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACACAGGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATAACCGACAAACCTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACATA 2074

RESULT 30
ADW53559/C
ID ADW53559 standard; cDNA; 6857 BP.
XX
AC ADW53559;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human 98P4B6 cDNA variant 25 SeqID 50.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytostatic; gene; chromosome 7q21.
XX
OS Homo sapiens.
XX
FN WO2003087306-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PR 06-JUN-2002; 2002US-00165044.
XX
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
XX WPI; 2003-903158/82.
XX
DR P-PSDB; ADW53560, ADW53596.
XX
XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
XX Example 2; SEQ ID NO 50; 616pp; English.

CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
CC 2, six transmembrane epithelial antigen of prostate-2) located on
CC chromosome 7q21 and the encoded protein and variants derived thereof.
CC Specifically, it refers to the expression pattern of this gene in adult
CC normal tissues and its aberrant over-expression in various cancers
CC including breast, colon, lung, kidney and prostate. The present invention
CC describes compositions and methods useful for detecting, preventing,
CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.
XX
SQ Sequence 6857 BP; 2019 A; 1318 C; 1209 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTGAAATTAACAGCTAACTCCCAAATGTACTTAACCTAGGGGAG 100
DB 2234 TTGAATATTCTTCATTGAAATTAACAGCTAACTCCCAAATGTCTGCTGTTTGCATAATGTTTAAACATTATATC 2175
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACCTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACACAGGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATAACCGACAAACCTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACATA 2074

RESULT 31
ADW53563/C
ID ADW53563 standard; cDNA; 6857 BP.
XX
AC ADW53563;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human 98P4B6 cDNA variant 27 SeqID 54.
XX
KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytostatic; gene; chromosome 7q21.
XX
OS Homo sapiens.
XX
FN WO2003087306-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010462.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PR 06-JUN-2002; 2002US-00165044.
XX
PR 20-DEC-2002; 2002US-0435480P.
XX
PA (AGEN-) AGENYSYS INC.
XX
PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX
XX WPI; 2003-903158/82.
XX
DR P-PSDB; ADW53564.
XX
XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

Example 2; SEQ ID NO 54; 616pp; English.

This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof. Specifically, it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2018 A; 1317 C; 1210 G; 2312 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

41 TTGAATATCTTCAATTTGAATAAACAAGCTTAACCTCCCAATTTGACCTAAGTAGGGGAG 100
|||||
2234 TTGAATATCTTCAATTTGAATAAACAAGCTTAAATTTGCTGGTTTGCATAATTTGTTTAAACATTATATC 2175
|||||
101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
|||||
2174 TATATTTAACTGTAAGTATTTTTCCTGCTGATTATGAATAATATAAGTAACAACAGGGTTA 2115
|||||
160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
|||||
2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2074
|||||

RESULT 32
ADW53525/c
ID ADW53525 standard; cDNA; 6857 BP.
XX ADW53525;
AC
XX 24-MAR-2005 (first entry)
XX Human 98P4B6 cDNA variant 8 SeqID 16.
XX
DE ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytostatic; gene; chromosome 7q21.
XX Homo sapiens.
OS
XX WO2003087306-A2.
PN
XX 23-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010462.
XX
XX 05-APR-2002; 2002US-0370387P.
PR
XX 06-JUN-2002; 2002US-00165044.
PR
XX 20-DEC-2002; 2002US-0435480P.
XX
XX (AGEN-) AGENYSYS INC.
PA
XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
PI
XX WPI; 2003-903158/82.
DR
XX P-PSDB; ADW53526, ADW53592.
XX
XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
PT detecting, preventing, prognosing and/or treating cancers that express

Example 2; SEQ ID NO 16; 616pp; English.

This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof. Specifically, it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

41 TTGAATATCTTCAATTTGAATAAACAAGCTTAACCTCCCAATTTGACCTAAGTAGGGGAG 100
|||||
2234 TTGAATATCTTCAATTTGAATAAACAAGCTTAAATTTGCTGGTTTGCATAATTTGTTTAAACATTATATC 2175
|||||
101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
|||||
2174 TATATTTAACTGTAAGTATTTTTCCTGCTGATTATGAATAATATAAGTAACAACAGGGTTA 2115
|||||
160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
|||||
2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2074
|||||

RESULT 33
ADW53565/c
ID ADW53565 standard; cDNA; 6857 BP.
XX ADW53565;
AC
XX 24-MAR-2005 (first entry)
XX Human 98P4B6 cDNA variant 28 SeqID 56.
XX
DE ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytostatic; gene; chromosome 7q21.
XX Homo sapiens.
OS
XX WO2003087306-A2.
PN
XX 23-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010462.
XX
XX 05-APR-2002; 2002US-0370387P.
PR
XX 06-JUN-2002; 2002US-00165044.
PR
XX 20-DEC-2002; 2002US-0435480P.
XX
XX (AGEN-) AGENYSYS INC.
PA
XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
PI
XX WPI; 2003-903158/82.
DR
XX P-PSDB; ADW53566.
XX

PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
PT detecting, preventing, prognosing and/or treating cancers that express
PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

DR	P-PSDB; ADW53568.
XX	
PT	A composition comprising 98P4B6 proteins and nucleic acid molecules for
PT	detecting, preventing, prognosing and/or treating cancers that express
PT	98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX	
XX	Example 2; SEQ ID NO 58; 616pp; English.
FS	
XX	
CC	This invention relates to a novel human gene 98P4B6 (also known as STEAP-
CC	2, six transmembrane epithelial antigen of prostate-2) located on
CC	chromosome 7q21 and the encoded protein and variants derived thereof.
CC	Specifically it refers to the expression pattern of this gene in adult
CC	normal tissues and its aberrant over-expression in various cancers
CC	including breast, colon, lung, kidney and prostate. The present invention
CC	describes compositions and methods useful for detecting, preventing,
CC	prognosing and/ or treating cancers that express 98P4B6. Furthermore, the
CC	polynucleotide can be used for gene therapy purposes, for monitoring
CC	genetic abnormalities and for generating transgenic or 'knockout',
CC	animal, which can be useful for the development and screening of
CC	therapeutically useful reagents. The encoded proteins may also be used in
CC	generating antibodies and vaccines, as well as in identifying ligands and
CC	cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC	these compositions exhibit cytostatic activities. This polynucleotide
CC	sequence is a human 98P4B6 cDNA variant of the invention.
CC	

Query Match	12.3%	Score 37	DB 11	Length 6857
Best Local Similarity	55.9%	Pred. No. 5.1		
Matches 90	Conservative 0	Mismatches 70	Indels 1	Gaps 1
Qy	41	TTGAAATATTCTTCATTTTGAATAAAACAGCTTAACCTCCCAAATTCCTAACTAGGGGAG	100	
Db	2234	TTGAAATATTTTTAAAGCTTTAAATTCCTGCTGGTTTGCAAATGTTTTTAACTATATC	2175	
Qy	101	TTTATCATGAAGAAATGTTTTTAAATGCTTTAAATTTTCAAATAAGAAAGTAACACAGAGTGTG	159	
Db	2174	TATATTATTAACTGTAAGTATTTTTGCTGCTGATTATGAATATATAAGTAAACAACAGGGTTA	2115	
Qy	160	CAACATGCTGTAAATAAACCCGACAAACTCTCAATCACTATA	200	
Db	2114	CAGTTGTGCAAAACCAAAACAAAAACATACCTTAAATTTACAATA	2074	

RESULT 34	
ADW53567/c	
ID: ADW53567 standard; cDNA; 6857 bp.	
XX:	
ADW53567;	
XX:	
AC	
XX:	
24-MAR-2005 (first entry)	
XX:	
Human 98P4B6 cDNA variant 29 SeqID 58.	
XX:	
ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;	
KW	
KW	
KW	
KW	
XX:	
XX:	
OS	
XX:	
WO2003087306-A2.	
PN	
XX:	
XX:	
23-OCT-2003.	
PD	
XX:	
XX:	
04-APR-2003; 2003WO-US010462.	
XX:	
05-APR-2002; 2002US-0370387P.	
PR	
06-JUN-2002; 2002US-00165044.	
PR	
20-DEC-2002; 2002US-0435480P.	
XX:	
XX:	
(AGEN-) AGENSYS INC.	
PA	
XX:	
Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;	
PI	
XX:	
WPI; 2003-903158/82.	
DR	

	Query Match	12.3%	Score 37	DB 11	Length 6857
	Best Local Similarity	55.9%	Pred. No. 5.1		
	Matches 90	Conservative 0	Mismatches 70	Indels 1	Gaps 1
Qy	41	TTGAAATATCTTCATTTGGAAATAAACAGCTAACTCCCAAAATGTACTTAACATAGGGGAG	100		
Db	2234	TTGAAATATTTTAAAGCTTAAATTCGCTGTTGCAAAATGTTTTTAAACATTATATC	2175		
Qy	101	TTTATCATGAA-GAAATGTTTTAAATGTTTAAATTAAGAAGTAAACAGAGTGTG	159		
Db	2174	TATATTATTAACGTGAAGTATTTTTTGCTCGATTATGAAATATAAGTAAACAAGGGTTA	2115		
Qy	160	CAACATGCTGTTAAATAACCCGACAAACTTCAATCACTATA	200		
Db	2114	CAGTTGTGCAAAACAAAAACAATACCTTAAATTTACAAATA	2074		

RESULT 35	
ADRR66608/c	
ID	ADRR66608 standard; DNA; 6857 BP.
XX	
AC	ADRR66608;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Human prostatic carcinoma derived DNA SEQ ID 120 #3.
XX	
KW	human; cytostatic; diagnosis; prostatic cancer;
XX	differential expression analysis; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO2004076614-A2.
XX	
PD	10-SEP-2004.
XX	
PF	22-FEB-2004; 2004WO-DE000433.
XX	
PR	27-FEB-2003; 2003DE-01009985.
PR	14-MAY-2003; 2003DE-01022134.
XX	
PA	(HINZ/) HINZMANN B.
PA	(DAHL/) DAHL E.
PA	(ROSE/) ROSENTHAL A.
PA	(HERM/) HERMANN K.
PA	(PILA/) PILARSKY C.

XX
PI Hinzmänn B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhang L, Staub E;
XX
XX MPI; 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
XX useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
XX Claim 1; Page 984-986; 1607pp; German.
XX
XX This invention describes novel cytostatic polynucleotide and polypeptide
XX sequences which can be used in a method for diagnosing prostatic cancer
XX or the risk of developing prostatic cancer. Diagnosis is based on
XX determining over transcription or over expression of the sequences in
XX prostatic tissue. Screening for inhibitors of the sequences or detection
XX substances involves a binding assay, any compounds that bind are
XX selected, optionally after deconvolution of mixtures. Detection of a
XX predetermined minimum level of the reporter indicates the presence of
XX tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
XX short-interfering RNA or ribozymes; an organic molecule of molecular
XX weight below 5000, preferably 300, that binds to the polypeptide; an
XX aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
XX polypeptide, preferably humanised or human; an anti-idiotype, non-human
XX (monoclonal) antibody directed against Ab or any of the above derivatised
XX with a reporter group, cell toxin, immunostimulatory molecules and/or
XX radioisotope. The polynucleotides are identified in human prostatic
XX cancer by differential expression analysis, using DNA microarrays,
XX between normal and tumorous tissues, with (over)expression being detected
XX by quantitative PCR. Analysis of prostatic cancer samples showed that
XX CD24 was upregulated in many of them. Sections of tissue, isolated from
XX prostatic cancer patients, or subjects at risk, were incubated
XX sequentially with anti-human CD4 murine monoclonal antibodies;
XX biotinylated second antibody; streptavidin-conjugated horseradish
XX peroxidase and then diaminobenzidine as colour former (brown). The
XX samples were counterstained with hemalum (blue). Malignant cells stained
XX strongly but non-malignant cells only weakly. In 15 of 63 samples of
XX adenocarcinoma, membrane and cytoplasmic staining was very strong, and
XX lymph node metastases were also stained. ADR65805-ADR66954 represent the
XX polynucleotide and polypeptide sequences used in the method of the
XX invention.
XX
XX Sequence 6857 BP; 2021 A; 1317 C; 1208 G; 2311 T; 0 U; 0 Other;
XX
XX Query Match 12.3%; Score 37; DB 13; Length 6857;
XX Best Local Similarity 55.9%; Pred. No. 5.1;
XX Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
XX
XX 41 TTGAATATTTCTTCATTGTAATAACAGCTAACTCCCAATTTCTACTTAAGGAG 100
XX 2234 TTGAATATTTTAAAGCTTAAATTCGTGTTTGCATTTGTAATTTTAACTATATC 2175
XX
XX 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
XX 2174 TATATTTATTAAGTCTGTAAGTATTTTTCCTGATTGTAATATAAGTAACACAGGTTA 2115
XX
XX 160 CAACATGCTGTTTAATTAACCCGACAACTTCAATCACTATA 200
XX 2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACAATA 2074
XX
XX
XX RESULT 36
XX ADR66266/C
XX ID ADR66266 standard; DNA; 6857 BP.
XX
XX ADR66266;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human prostatic carcinoma derived DNA SEQ ID 120 #2.
XX

KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
XX Homo sapiens.
XX
XX WO2004076614-A2.
XX
XX 10-SEP-2004.
XX
XX 22-FEB-2004; 2004WO-DE000433.
XX
XX 27-FEB-2003; 2003DE-01009985.
XX 14-MAY-2003; 2003DE-01022134.
XX
XX (HINZ/) HINZMANN B.
XX (DAHL/) DAHL E.
XX (ROSE/) ROSENTHAL A.
XX (HERM/) HERMANN K.
XX (PILA/) PILARSKY C.
XX
XX Hinzmänn B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
XX Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
XX Xinzhang L, Staub E;
XX
XX MPI; 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
XX useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
XX Claim 1; Page 338-339; 1607pp; German.
XX
XX This invention describes novel cytostatic polynucleotide and polypeptide
XX sequences which can be used in a method for diagnosing prostatic cancer
XX or the risk of developing prostatic cancer. Diagnosis is based on
XX determining over transcription or over expression of the sequences in
XX prostatic tissue. Screening for inhibitors of the sequences or detection
XX substances involves a binding assay, any compounds that bind are
XX selected, optionally after deconvolution of mixtures. Detection of a
XX predetermined minimum level of the reporter indicates the presence of
XX tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
XX short-interfering RNA or ribozymes; an organic molecule of molecular
XX weight below 5000, preferably 300, that binds to the polypeptide; an
XX aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
XX polypeptide, preferably humanised or human; an anti-idiotype, non-human
XX (monoclonal) antibody directed against Ab or any of the above derivatised
XX with a reporter group, cell toxin, immunostimulatory molecules and/or
XX radioisotope. The polynucleotides are identified in human prostatic
XX cancer by differential expression analysis, using DNA microarrays,
XX between normal and tumorous tissues, with (over)expression being detected
XX by quantitative PCR. Analysis of prostatic cancer samples showed that
XX CD24 was upregulated in many of them. Sections of tissue, isolated from
XX prostatic cancer patients, or subjects at risk, were incubated
XX sequentially with anti-human CD4 murine monoclonal antibodies;
XX biotinylated second antibody; streptavidin-conjugated horseradish
XX peroxidase and then diaminobenzidine as colour former (brown). The
XX samples were counterstained with hemalum (blue). Malignant cells stained
XX strongly but non-malignant cells only weakly. In 15 of 63 samples of
XX adenocarcinoma, membrane and cytoplasmic staining was very strong, and
XX lymph node metastases were also stained. ADR65805-ADR66954 represent the
XX polynucleotide and polypeptide sequences used in the method of the
XX invention.
XX
XX Sequence 6857 BP; 2021 A; 1317 C; 1208 G; 2311 T; 0 U; 0 Other;
XX
XX Query Match 12.3%; Score 37; DB 13; Length 6857;
XX Best Local Similarity 55.9%; Pred. No. 5.1;
XX Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
XX
XX 41 TTGAATATTTCTTCATTGTAATAACAGCTAACTCCCAATTTCTACTTAAGGAG 100
XX 2234 TTGAATATTTTAAAGCTTAAATTCGTGTTTGCATTTGTAATTTTAACTATATC 2175
XX

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACACAGAGTTG 159
 DB 2174 TATATTATTAACGTGAAGTATTTTGGCTGATTATGAATTAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATACCGACAAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2074

RESULT 37
 AEC20351/c
 ID: AEC20351 standard; cDNA; 6857 BP.
 XX
 AC AEC20351;
 XX
 DT 03-NOV-2005 (first entry)
 XX
 DE cDNA encoding human STEAP2.
 XX
 KW pharmaceutical; six transmembrane epithelial antigen of prostate 2;
 KW STEAP2; antibody; cancer; neoplasm; cell growth; ADEPT; colon tumor;
 KW breast tumor; lung tumor; ovary tumor; pancreas tumor; expression;
 KW diagnosis; solid tumor; hyperproliferation; immune disorder;
 KW autoimmune disease; allergy; inflammation; cytostatic; antipsoriatic;
 KW dermatological; antiarthritic; immunosuppressive; antiallergic;
 KW antiinflammatory; vaccine; gene; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 394..1866
 /*tag= a
 /product= "STEAP2"
 WO2005079490-A2.
 01-SEP-2005.
 11-FEB-2005; 2005WO-US005204.
 13-FEB-2004; 2004US-00779460.
 (NUVE-) NUVELO INC.
 Entage PCR;
 WPI: 2005-591919/60.
 P-PSDB; AEC20352.
 Treating cancer comprises use of an anti-STEAP2 (six transmembrane epithelial antigen of prostate 2) antibody, or other agent targeting STEAP2.

Claim 15; SEQ ID NO 1; 96pp; English.
 The invention relates to a pharmaceutical composition comprising an anti-six transmembrane epithelial antigen of prostate 2 (STEAP2) antibody specific for cancer cells, where the antibody specifically binds to a STEAP2 polypeptide having a fully defined sequence given as SEQ ID No:2 in the specification. Also described are: (1) a method (M1) of killing or inhibiting the growth of STEAP2-expressing cells by contacting the cells with a composition comprising: (a) an antibody specific for a polypeptide of SEQ ID NO:2, (b) a conjugate comprising an antibody linked to a prodrug-activating enzyme, and a prodrug convertible by the conjugate into a cytotoxic drug, (c) an anti-STEAP2 antigen, (d) a vector comprising a nucleic acid encoding STEAP2 (SEQ ID NO:1), or its immunogenic fragment, (e) an antigen-presenting cell comprising the vector, (f) a small molecule specifically binding to a polypeptide of SEQ ID NO:2 or its immunogenic fragment, (g) a polypeptide specifically binding to a polypeptide of SEQ ID NO:2 or its immunogenic fragment; (2) a method (M2) of diagnosing a cancer of the colon, breast, lung, ovary or pancreas, comprising: (a) detecting or measuring expression of STEAP2 by the cancer, and (b) comparing the expression to a standard indicative of

CC the cancer and/or to normal tissue; (3) a method (M3) for the use of any of the agents listed in (1) in the preparation of a medicament for killing or inhibiting the growth of STEAP2-expressing cells of a cancer of the colon, breast, lung, ovary or pancreas. The composition of the invention and method (M1) are useful for treating cancer, especially cancer of the colon, breast, lung, ovary or pancreas. Method (M2) is useful for diagnosing any of these cancers. Agents targeting STEAP2 are also disclosed as being potentially useful for treating other solid tumors and hematopoietic-based cancers and other hyperproliferative disorders such as X-linked lymphoproliferative disorders, Epstein-Barr virus-related conditions such as mononucleosis, hyperplasia, psoriasis, contact dermatitis and immunological disorders, arthritis, autoimmune disease, allergy and inflammation. This sequence encodes human STEAP2.
 CC Note: The sequence data for this patent was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 6857 BP; 2021 A; 1317 C; 1208 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 14; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 41 TTCAATATTTCTTCATTGAAATAAACAGCTAACTCCCAATTTGTACTAACTAGGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCAATTTGTTTAACTATTATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAGAGTAACACAGAGTTG 159
 DB 2174 TATATTATTAACGTGAAGTATTTTGGCTGATTATGAATATTAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATACCGACAAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2074

RESULT 38
 AAH18058
 ID AAH18058 standard; cDNA; 2272 BP.
 XX
 AC AAH18058;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:17892.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EF1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

10

PR	16-MAR-2000;	2000US-01898744P
PR	17-MAR-2000;	2000US-01900763P
PR	18-APR-2000;	2000US-01981237P
PR	19-MAY-2000;	2000US-02055159P
PR	07-JUN-2000;	2000US-0209467P
PR	28-JUN-2000;	2000US-0214886P
PR	30-JUN-2000;	2000US-0215835P
PR	07-JUL-2000;	2000US-0216647P
PR	11-JUL-2000;	2000US-0216880P
PR	11-JUL-2000;	2000US-0217487P
PR	14-JUL-2000;	2000US-0217496P
PR	14-JUL-2000;	2000US-0218290P
PR	26-JUL-2000;	2000US-0220963P
PR	26-JUL-2000;	2000US-0220964P
PR	26-JUL-2000;	2000US-0224518P
PR	14-AUG-2000;	2000US-0224519P
PR	14-AUG-2000;	2000US-0225213P
PR	14-AUG-2000;	2000US-0225214P
PR	14-AUG-2000;	2000US-0225266P
PR	14-AUG-2000;	2000US-0225267P
PR	14-AUG-2000;	2000US-0225268P
PR	14-AUG-2000;	2000US-0225270P
PR	14-AUG-2000;	2000US-0225447P
PR	14-AUG-2000;	2000US-0225757P
PR	14-AUG-2000;	2000US-0225758P
PR	14-AUG-2000;	2000US-0225759P
PR	18-AUG-2000;	2000US-0226279P
PR	22-AUG-2000;	2000US-0226681P
PR	22-AUG-2000;	2000US-0226868P
PR	22-AUG-2000;	2000US-0227182P
PR	30-AUG-2000;	2000US-0227009P
PR	01-SEP-2000;	2000US-0228924P
PR	01-SEP-2000;	2000US-0229287P
PR	01-SEP-2000;	2000US-0229343P
PR	01-SEP-2000;	2000US-0229344P
PR	01-SEP-2000;	2000US-0229345P
PR	05-SEP-2000;	2000US-0229509P
PR	06-SEP-2000;	2000US-0229513P
PR	08-SEP-2000;	2000US-0230437P
PR	08-SEP-2000;	2000US-0230438P
PR	08-SEP-2000;	2000US-0231242P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231244P
PR	08-SEP-2000;	2000US-0231413P
PR	08-SEP-2000;	2000US-0231414P
PR	08-SEP-2000;	2000US-0232080P
PR	08-SEP-2000;	2000US-0232081P
PR	12-SEP-2000;	2000US-0231968P
PR	14-SEP-2000;	2000US-0232397P
PR	14-SEP-2000;	2000US-0232398P
PR	14-SEP-2000;	2000US-0232399P
PR	14-SEP-2000;	2000US-0232400P
PR	14-SEP-2000;	2000US-0234011P
PR	14-SEP-2000;	2000US-0233063P
PR	14-SEP-2000;	2000US-0233064P
PR	14-SEP-2000;	2000US-0233065P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234274P
PR	25-SEP-2000;	2000US-0234597P
PR	25-SEP-2000;	2000US-0234988P
PR	26-SEP-2000;	2000US-0235484P
PR	27-SEP-2000;	2000US-0235834P
PR	27-SEP-2000;	2000US-0235836P
PR	29-SEP-2000;	2000US-0236327P
PR	29-SEP-2000;	2000US-0236367P
PR	29-SEP-2000;	2000US-0236368P
PR	29-SEP-2000;	2000US-0236369P
PR	29-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-0236802P
PR	02-OCT-2000;	2000US-0237037P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237040P

XX PI Lesche R, Fassbender A, Juenemann K, Foekens J, Martens JMW;
 XX DR WPI; 2006-125479/13.
 XX PT Detecting or detecting and distinguishing between or among breast cell
 PT Proliferative disorders in a subject by determining the expression of a
 XX panel of genes or sequences.
 XX PS Claim 14; SEQ ID NO 913; 321pp; English.
 XX CC The invention relates to a method of detecting or detecting and
 CC distinguishing between or among breast cell proliferative disorders in a
 CC subject comprising determining the expression of at least one gene or
 CC sequence given in the specification, and determining from the expression
 CC the presence, absence or subclass of a breast cell proliferative
 CC disorder. The invention also relates to a treated nucleic acid used to
 CC convert at least one unmethylated cytosine base of the genomic DNA
 CC sequence to uracil or another base that is detectably dissimilar to
 CC cytosine in terms of hybridization, and an oligomer that is complementary
 CC to, or hybridizes under moderately stringent conditions to a treated
 CC genomic DNA sequence or a set of oligomers comprising at least two
 CC oligonucleotides. The method is useful for detecting or detecting and
 CC distinguishing between or among breast cell proliferative disorders in a
 CC subject. This sequence represents human genomic DNA used in the scope of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2374 BP; 536 A; 0 C; 675 G; 1163 T; 0 U; 0 Other;
 SQ
 Query Match 12.1%; Score 36.4; DB 15; Length 2374;
 Best Local Similarity 47.1%; Pred. No. 5.7;
 Matches 112; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 Qy 37 AGACTGGAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTACCTAAGTAGG 96
 Db 275 AAAATTTAAATATATTTTAAATTAACAAACACCTTAAATTTAAATTAACCTCAATATAT 216
 Qy 97 GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACCCAGAGTG 156
 Db 215 CACTTTTTTAAATACATATATTTAAACCTTAAATTAAGAAAGTAACCTTAAATTTTAAAT 156
 Qy 157 TTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCA 216
 Db 155 AAATAAAATCATATACTATAAACAACCTTAAATTAAGAAAGTAACCTTAAATTAAGAAAGCA 96
 Qy 217 TTCTGCAAGGATCCAGAGTAACAGATATTTTGAATGCAATGTTGAACCGACCATATA 274
 Db 95 ACACACAAACTATAATAAACAATAATTAATAAATACTATATTTATTACTAACTCTA 38
 RESULT 45
 AEF55696/c
 ID AEF55696 standard; DNA; 2374 BP.
 XX AC AEF55696;
 XX AC AEF55696;
 XX DT 23-MAR-2006 (first entry)
 XX DE Human genomic DNA #303.
 XX KW Diagnosis; breast disease; hyperproliferation; gene; ds.
 XX OS Homo sapiens.
 XX FN WO2006008128-A2.
 XX PD 26-JAN-2006.
 XX PF 18-JUL-2005; 2005WO-EP007830.
 XX PR 18-JUL-2004; 2004EP-00016926.

XX PA (EPIG-) EPIGENOMICS AG.
 XX PI Lesche R, Fassbender A, Juenemann K, Foekens J, Martens JMW;
 XX DR WPI; 2006-125479/13.
 XX PT Detecting or detecting and distinguishing between or among breast cell
 PT Proliferative disorders in a subject by determining the expression of a
 XX panel of genes or sequences.
 XX PS Claim 14; SEQ ID NO 677; 321pp; English.
 XX CC The invention relates to a method of detecting or detecting and
 CC distinguishing between or among breast cell proliferative disorders in a
 CC subject comprising determining the expression of at least one gene or
 CC sequence given in the specification, and determining from the expression
 CC the presence, absence or subclass of a breast cell proliferative
 CC disorder. The invention also relates to a treated nucleic acid used to
 CC convert at least one unmethylated cytosine base of the genomic DNA
 CC sequence to uracil or another base that is detectably dissimilar to
 CC cytosine in terms of hybridization, and an oligomer that is complementary
 CC to, or hybridizes under moderately stringent conditions to a treated
 CC genomic DNA sequence or a set of oligomers comprising at least two
 CC oligonucleotides. The method is useful for detecting or detecting and
 CC distinguishing between or among breast cell proliferative disorders in a
 CC subject. This sequence represents human genomic DNA used in the scope of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2374 BP; 536 A; 146 C; 675 G; 1017 T; 0 U; 0 Other;
 SQ
 Query Match 12.1%; Score 36.4; DB 15; Length 2374;
 Best Local Similarity 47.1%; Pred. No. 5.7;
 Matches 112; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 Qy 37 AGACTGGAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTACCTAAGTAGG 96
 Db 275 AAAATTTAAATATATTTTAAATTAACAAACACCTTAAATTTAAATTAACCTCAATATAT 216
 Qy 97 GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACCCAGAGTG 156
 Db 215 CACTTTTTTAAATACATATATTTAAACCTTAAATTAAGAAAGTAACCTTAAATTTTAAAT 156
 Qy 157 TTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCA 216
 Db 155 AAATAAAATCATATACTATAAACAACCTTAAATTAAGAAAGTAACCTTAAATTAAGAAAGCA 96
 Qy 217 TTCTGCAAGGATCCAGAGTAACAGATATTTTGAATGCAATGTTGAACCGACCATATA 274
 Db 95 ACACACAAACTATAATAAACAATAATTAATAAATACTATATTTATTACTAACTCTA 38
 RESULT 46
 ACL35363
 ID ACL35363 standard; cDNA; 2000 BP.
 XX AC ACL35363;
 XX AC ACL35363;
 XX DT 02-JUN-2005 (first entry)
 XX DE Rice stress-regulated promoter SEQ ID NO:13926.
 XX KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX KW agriculture.
 XX OS Oryza sativa.
 XX FN WO2003008540-A2.
 XX PD 30-JAN-2003.
 XX PR

21-JUN-2002; 2002WO-US019668.
22-JUN-2001; 2001US-0300112P.
24-AUG-2001; 2001US-0314662P.
26-SEP-2001; 2001US-0325277P.
21-NOV-2001; 2001US-0332132P.
(SYGN) SYNGENTA PARTICIPATIONS AG.
Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri P,
Moughamer T, Provart N, Rieke D, Zhu T;
WPI; 2003-248011/24.
New stress-responsive nucleic acid, useful for altering the
responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
stress, salt stress or osmotic stress.
Claim 48; SEQ ID NO 13926; 89pp; English.
The invention relates to novel abiotic stress responsive polynucleotides
and polypeptides. Also disclosed are vectors, expression cassettes, host
cells, and plants containing such polynucleotides. Also disclosed are
methods for using the polynucleotides and polypeptides to alter the
responsiveness of a plant to abiotic stress. The invention is useful in
agriculture. The nucleic acid is useful for determining whether a test
plant has been exposed to an abiotic stress condition. It is also useful
for selecting an agent that alters abiotic stress regulated
polynucleotide expression in a plant cell, and to identify a homolog or
ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
molecule and the polypeptide encoded by it are useful in altering the
responsiveness of a plant to an abiotic stress, such as cold stress, salt
stress, osmotic stress or any of their combinations. The present sequence
is used in the exemplification of the invention
Sequence 2000 BP; 540 A; 344 C; 416 G; 428 T; 0 U; 272 Other;
Query Match 12.1%; Score 36.2; DB 11; Length 2000;
Best Local Similarity 12.3%; Pred. NO. 6.1;
Matches 29; Conservative 109; Mismatches 97; Indels 0; Gaps 0;
17 ATGCTAGTACTAGCAGAGCTGGAATATCTTCATTGAAATAAACAGCTAATC 76
123 WTRWMMWKACTCYTMYTKTMYTWRWAAAWMMWKYKMWARGWSAKMKAYRYRKCTY 182
77 CCAATTTGCTAGCTAGGAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCA 136
183 WAWRMAWRMTTKWCTWSWAWTCTYATWAWATATWTTTRWAWKTRAWMTWRYAW 242
137 AATAAGAGTAACACAGAGTGTTCGAACATGCTGTTAAATAACCCGACAAATCTCAATCAC 196
243 MYTWWAKKAWKYATKWTTYAWRMMWMSRTYRWMTWAWMYTWSYTYTWAWSY 302
197 TATAGCTGTAGTAGTGCATCTTCGAAGGATCCAGAGTACCAAGTATTTTGA 251
303 WWCMTWTSRRRWYRYSKRRRTWTTWTTWMTWMMKWKYKAWKWWMAWKW 357
RESULT 47
ABK31277/c
ID ABK31277 standard; DNA; 9087 BP.
XX ABK31277;
XX
XX 23-APR-2002 (first entry)
XX Signal transduction associated gene modified complementary DNA #60.
XX Human; signal transduction associated gene; cytosine methylation state;
XX CpG island; signal transduction associated disease; solid tumour; cancer;
XX antitumour; cytostatic; mutant; ds.
XX Homo sapiens.
OS

OS Synthetic.
XX WO200200926-A2.
XX
XX 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP007472.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-147896/19.
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
XX signal transduction e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with signal transduction.
XX Claim 1; SEQ ID NO 120; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
XX signal transduction associated genes. The DNA sequences are chemically
XX modified using a solution of bisulphite, hydrogen sulphite or disulphite.
XX Also disclosed are oligonucleotides and/or PNA oligomers for detecting
XX the cytosine methylation state (CpG islands) of these genes, and a method
XX for the diagnosis and/or therapy of genetic and epigenetic parameters of
XX genes associated with signal transduction. The genomic DNA can be
XX obtained from cells or cellular components which contain DNA, e.g. cell
XX lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
XX tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
XX brain, heart, prostate, lung, breast or liver, histologic object slides,
XX and all their possible combinations. The sequences of the invention are
XX useful for the diagnosis and therapy of diseases associated with signal
XX transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
XX chemically pretreated genomic DNA sequences of different genes associated
XX with signal transduction, or their complementary sequences. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office
XX
XX Sequence 9087 BP; 2207 A; 253 C; 2592 G; 4034 T; 0 U; 1 Other;
Query Match 12.1%; Score 36.2; DB 6; Length 9087;
Best Local Similarity 48.3%; Pred. No. 9.1; Mismatches 108; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 79 AAATTGTACCTTAACCTAGGGAGTGTTCATCATGAAGAAATGTTTAAATGCTTAAATTTTCAA 138
DB 1413 AAAATTCCTCTTAAAAAAACCTTACCTTAAATTAATAATAATATCTATTACTACT 1354
QY 139 TAAGAAGTAACGAGAGTGTTCGAACATGCTGTTTAAATAACCCGACAACTTCAATCACTA 198
DB 1353 CACAATAAAACAATAACTACAAAACATATTCGAAAAAACCTCTTAACAAAATTTCTCAATA 1294
QY 199 TAGCTGTAGTAGAGTGCATTTCTGCAAGATCCAGAGTAAACAGTATTTTGGAAATGCAA 258
DB 1293 AACCCCTACATATCTCTTCTCTCCAAACAACCTTAATCGAAAAATCATCTAAAAAACCCTAA 1234
QY 259 TGTGTAACCCGACCATACATACTAATTTATCTGCT 287
DB 1233 CCTTACCCCTCCCTTAAAAAACAATCTCT 1205
RESULT 48
ABL70238/c
ID ABL70238 standard; DNA; 9087 BP.
XX
XX ABL70238;
XX
XX 01-JUL-2002 (first entry)
DT

XX DE Chemically treated cell signalling DNA sequence complementary to#64.
XX DE Cell signalling; cytosine methylation; cell signalling disease; cancer;
XX DE tumour; cytostatic; ds.
XX KW Unidentified.
XX OS WO200202807-A2.
XX PN 10-JAN-2002.
XX PD 29-JUN-2001; 2001WO-EP007471.
XX PF 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX XX (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX XX WPI; 2002-154758/20.
XX XX Nucleic acid, useful for diagnosis and therapy of diseases associated
XX PT with cell signaling e.g. Cancer, comprises chemically modified genomic
XX PT sequences of genes associated with cell signaling.
XX XX Claim 1; SEQ ID NO 128; 24pp + Sequence Listing; English.
XX XX The invention relates to a nucleic acid comprising a sequence of at least
XX CC 18 bases of a segment of chemically pretreated DNA of genes associated
XX CC with cell signalling. The activity of the modified sequences of the
XX CC invention may be described as cytostatic. The object of the invention is
XX CC to provide the chemically modified DNA of genes associated with cell
XX CC signalling, as well as oligonucleotides and/or PNA-oligonucleotides for
XX CC detecting cytosine methylations, as well as a method which is
XX CC particularly suitable for the diagnosis and/or therapy of genetic and
XX CC epigenetic parameters of genes associated with cell signalling. The
XX CC chemically modified DNA provided by the invention is useful for diagnosis
XX CC and therapy of diseases such as solid tumours and cancer. The sequences
XX CC given in records ABL70111-ABL70626 represent chemically pre-treated
XX CC genomic DNA's of genes associated with cell signalling. Note: The
XX CC sequence data for this patent is not represented in the printed
XX CC specification, but is based on sequence information supplied by the
XX CC European Patent Office
XX SQ Sequence 9087 BP; 2207 A; 253 C; 2592 G; 4034 T; 0 U; 1 Other;
Query Match 12.1%; Score 36.2; DB 6; Length 9087;
Best Local Similarity 48.3%; Pred. No. 9.1;
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 79 AAATTGTACCTAACTAGGGAGTTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAA 138
DB 1413 AAAATTCCTTCTTAAAAAACCTTACCCTAAATTAATAATAATATCTATTACTACT 1354
QY 139 TAAGAAGTAACAGAGTGTTCGACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
DB 1353 CACAATAAACAATAACTACAAACATATTTCGCAAAAACCTCTTAACAAATTTCTCAATA 1294
QY 199 TAGCTGTAGTAGAGTGCAATTCGAGAGATCCGAGAGTACAGTATTTTGGAAATGCAA 258
DB 1293 AACCCCTACAATACTTCCTTCTCCAAACAACCTTAATCGAAATAATCTCTAAAAACCTAA 1234
QY 259 TGTGTGAACCGACCATACTAATTAATTCGTCT 287
DB 1233 CCTTACCCCTCCCTAAAAACAACCTTCT 1205
RESULT 49
AAS61181/c
ID AAS61181 standard; DNA; 9087 BP.
XX

AC AAS61181;
XX 29-JAN-2002 (first entry)
XX DE Human gene regulation-associated gene oligonucleotide #136.
XX XX
XX KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
XX KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
XX KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
XX KW renal disease; Preeclampsia; cardiac allograft vascular disease;
XX KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
XX KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
XX KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX OS Homo sapiens.
XX XX WO200177375-A2.
XX PN 18-OCT-2001.
XX PD 06-APR-2001; 2001WO-EP003968.
XX PF 06-APR-2000; 2000DE-01019058.
XX XX 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX XX (EPIG-) EPIGENOMICS AG.
XX PA Olek A, Piepenbrock C, Berlin K;
XX PI WPI; 2002-017470/02.
XX DR New nucleic acid sequences from chemically modified genes associated with
XX PT gene regulation, useful for analyzing cytosine methylations for diagnosis
XX PT and therapy of diseases e.g. severe combined immunodeficiency disease.
XX XX Claim 1; SEQ ID NO 140; 26pp; English.
XX PS The invention relates to 224 nucleic acid sequences comprising at least
XX CC 18 bases of a chemically pretreated gene associated with gene regulation
XX CC selected from 43 known genes (or complementary sequences). The chemical
XX CC pretreatment converts cytosine bases unmethylated at the 5-position to
XX CC uracil or another base with hybridisation behaviour dissimilar to
XX CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
XX CC oligomers (or sets/arrays) and method are useful in the diagnosis of
XX CC diseases (or predisposition to diseases) associated with gene regulation
XX CC and in therapy of such diseases, by enabling analysis of the cytosine
XX CC methylation patterns of such genes, kits are provided. They are
XX CC especially useful in diagnosis and therapy of e.g. severe combined
XX CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
XX CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
XX CC syndrome, renal disease, preeclampsia, graft versus-host disease. The
XX CC present sequence is a sequence included in the sequence data for this
XX CC specification and is associated with the human gene regulation-associated
XX CC genes. Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 9087 BP; 2207 A; 253 C; 2592 G; 4034 T; 0 U; 1 Other;
Query Match 12.1%; Score 36.2; DB 6; Length 9087;
Best Local Similarity 48.3%; Pred. No. 9.1;
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 79 AAATTGTACCTAACTAGGGAGTTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAA 138
DB 1413 AAAATTCCTTCTTAAAAAACCTTACCCTAAATTAATAATAATATCTATTACTACT 1354
QY 139 TAAGAAGTAACAGAGTGTTCGACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
DB 1353 CACAATAAACAATAACTACAAACATATTTCGCAAAAACCTCTTAACAAATTTCTCAATA 1294

Oy 199 TAGCTGTAGTAGTGCATTCTGCAAGGATCCGAGTAGTAACAGTATTTTGGAAATGCAA 258
Db 1293 AACCCCTACAATCTCTCTTCCAAACAATTAATCGAAAAATCATCTAAAAACCTAA 1234
Oy 259 TGTTGAACCGACCATACTAATTTATCTGCT 287
Db 1233 CCYTACCCCTCCCTTAATAAAACAACCTTCT 1205

RESULT 50

ABL33538/C

ID ABL33538 standard; DNA; 17738 BP.

AC ABL33538;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1511.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antianaemic; cytotatic; nootropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.

XX Claim 1; SEQ ID NO 1511; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention

XX Sequence 17738 BP; 4696 A; 289 C; 3655 G; 9098 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 12.1%; Score 36.2; DB 6; Length 17738;

XX Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Oy 40 CTTGAAATATTTCTTCAATTAATAACAGCTAACTCCCAATTTGTACCTAACTAGGGGA 99
Db 9292 CTTATCTTATTTCAATCTTAAAAAAAATCTTTCTTAATTTATACACAATACATA 9233

Oy 100 GTTTATCATGAGAAATGTTTAAATGCTTATTTTCAATTAAGTAGTACCAAGTGTG 159

Db 9232 TAATCTTATCTATCAAAAAAAATAAATAATTTAAAAAAAATTCAAAAAAATTTTA 9173

Oy 160 CAACATGCTGTTAATAAATACCCGACAAACTTCAATCACTATA 200
Db 9172 AAAAAATTATATAAATTTATCTCAAAAACCTTAAAAAACTAAA 9132

Search completed: July 17, 2006, 21:54:57
Job time : 396 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:55:10 ; Search time 2892 Seconds
(without alignments)
5800.761 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1 ggcattgacggctaaatgc.....atctgctgattaggaagstat 300

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST.*

1: gb_est1.*

2: gb_est3.*

3: gb_est4.*

4: gb_est5.*

5: gb_est6.*

6: gb_est2.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_gss1.*

11: gb_gss2.*

12: gb_gss3.*

13: gb_gss4.*

14: gb_gss4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141.4	29.1	715	13	CW889980
2	88.2	29.4	183	13	DU405174
3	45.8	15.3	684	2	BJ924615
4	45.8	15.3	765	2	BJ919954
5	43.6	14.5	411	8	C0338585
6	43.6	14.5	514	8	C0275038
7	43.6	14.5	545	8	C0264525
8	43.6	14.5	604	8	C0312735
9	43	14.3	818	11	AZ123198
10	42	14.0	399	8	C0185341
11	42	14.0	1933	6	AK038018
12	41.4	13.8	684	11	BZ281627
13	40.8	13.6	674	9	DN248758
14	40.4	13.5	788	11	BH463846
15	40.2	13.4	694	13	CW144136
16	40.2	13.4	1172	12	CC272514
17	39.6	13.2	543	5	CD288289
18	39.6	13.2	712	5	CJ361302
19	39.6	13.2	743	14	CT404992

20	39.6	13.2	745	5	CJ381709
21	39.6	13.2	810	5	CJ385948
22	39.4	13.1	539	13	CZ012938
23	39.4	13.1	758	14	DU930095
24	38.8	12.9	547	12	CG804676
25	38.8	12.9	690	11	BH929322
26	38.6	12.9	359	7	BF336378
27	38.6	12.9	759	11	BZ049137
28	38.4	12.8	535	14	CR068462
29	38.4	12.8	683	13	CL341203
30	38.4	12.8	746	9	DN376396
31	38.4	12.8	779	9	CX973570
32	38.2	12.7	756	8	CV975922
33	38.2	12.7	759	11	BH508774
34	38.2	12.7	808	13	CZ764379
35	38.2	12.7	818	13	DU208665
36	38.2	12.7	3792	11	AF156725
37	38	12.7	559	11	AO836871
38	38	12.7	1091	10	DM602832
39	37.8	12.6	914	13	CZ953670
40	37.8	12.6	950	14	DU931319
41	37.6	12.5	684	4	CB430880
42	37.6	12.5	714	11	AO855958
43	37.6	12.5	837	11	BH486214
44	37.6	12.5	991	3	BU355289
45	37.4	12.5	373	10	R22138
46	37.4	12.5	514	11	AQ321454
47	37.4	12.5	563	11	AQ627491
48	37.4	12.5	704	13	CW685077
49	37.4	12.5	723	11	BZ047105
50	37.4	12.5	797	11	BZ479314
51	37.4	12.5	896	13	CZ529782
52	37.4	12.5	940	13	DU080282
53	37.2	12.4	416	8	CO343761
54	37.2	12.4	486	7	BE856330
55	37.2	12.4	521	1	AV606314
56	37.2	12.4	551	1	AV602756
57	37.2	12.4	647	4	CB446831
58	37.2	12.4	660	4	CB425397
59	37.2	12.4	672	8	CN440384
60	37.2	12.4	683	4	CB464131
61	37.2	12.4	693	8	CV974819
62	37.2	12.4	700	8	CO877917
63	37.2	12.4	788	9	CX713951
64	37.2	12.4	795	9	DN545081
65	37.2	12.4	830	9	DN543711
66	37.2	12.4	1080	1	AJ925742
67	37.2	12.4	1253	10	DV779655
68	37	12.3	366	5	CK779724
69	37	12.3	527	4	EX490203
70	37	12.3	613	4	EX641441
71	37	12.3	618	11	AQ244513
72	37	12.3	621	8	CN762630
73	37	12.3	628	13	CW144135
74	37	12.3	673	4	EX504981
75	37	12.3	707	4	EX509011
76	37	12.3	733	14	AG037123
77	37	12.3	778	13	CZ465308
78	37	12.3	805	13	DU209109
79	37	12.3	809	14	CT185225
80	37	12.3	884	14	DU888537
81	37	12.3	906	3	BQ948933
82	37	12.3	945	14	CNS04D0K
83	36.8	12.3	301	2	BG461576
84	36.8	12.3	307	2	BG461289
85	36.8	12.3	368	8	CN360568
86	36.8	12.3	498	3	BQ024056
87	36.8	12.3	558	8	CV221846
88	36.8	12.3	566	3	BU752970
89	36.8	12.3	632	7	BE631244
90	36.8	12.3	642	3	BN985293
91	36.8	12.3	676	3	BN683514
92	36.8	12.3	680	4	CA056323

CJ381709	CJ381709
CJ385948	CJ385948
CZ012938	CZ012938
DU930095	DU930095
CG804676	CG804676
BH929322	BH929322
BF336378	BF336378
BZ049137	BZ049137
CR068462	CR068462
CL341203	CL341203
DN376396	DN376396
CX973570	CX973570
CV975922	CV975922
BH508774	BH508774
CZ764379	CZ764379
DU208665	DU208665
AF156725	AF156725
AO836871	AO836871
DM602832	DM602832
CZ953670	CZ953670
DU931319	DU931319
CB430880	CB430880
AO855958	AO855958
BH486214	BH486214
BU355289	BU355289
R22138	R22138
AQ321454	AQ321454
AQ627491	AQ627491
CW685077	CW685077
BZ047105	BZ047105
BZ479314	BZ479314
CZ529782	CZ529782
DU080282	DU080282
CO343761	CO343761
BE856330	BE856330
AV606314	AV606314
AV602756	AV602756
CB446831	CB446831
CB425397	CB425397
CN440384	CN440384
CB464131	CB464131
CV974819	CV974819
CO877917	CO877917
CX713951	CX713951
DN545081	DN545081
DN543711	DN543711
AJ925742	AJ925742
DV779655	DV779655
CK779724	CK779724
EX490203	EX490203
EX641441	EX641441
AQ244513	AQ244513
CN762630	CN762630
CW144135	CW144135
EX504981	EX504981
EX509011	EX509011
AG037123	AG037123
CZ465308	CZ465308
DU209109	DU209109
CT185225	CT185225
DU888537	DU888537
BQ948933	BQ948933
CNS04D0K	CNS04D0K
BG461576	BG461576
BG461289	BG461289
CN360568	CN360568
BQ024056	BQ024056
CV221846	CV221846
BU752970	BU752970
BE631244	BE631244
BN985293	BN985293
BN683514	BN683514
CA056323	CA056323

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.

REFERENCE 1 (bases 1 to 183)

AUTHORS Kirkness, E., Shetty, J., de Jong, P., McEwan, J.C., Oddy, H. and Cockett, N.

TITLE Ovine BAC End Sequences from Library CHORI-243

JOURNAL Unpublished (2004)

COMMENT Other GSSs: 1098473751940

Contact: Ewen Kirkness

The Institute for Genomic Research (TIGR; www.tigr.org)

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-795-7536

Email: ekirknes@tigr.org

Sequences generated at the J. Craig Venter Institute Joint

Technology Center (JCVITC; <http://www.venterlinstitute.org/>).

Original Trace: 1098448345721 Trace TI: gml.ti.918949075

Insert Length: 184000 Std Error: 0.00 row: B column: 8

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .183

/organism="Ovis aries"

/mol_type="genomic DNA"

/strain="Texel breed"

/db_xref="taxon:9940"

/clones="CH243-193B8"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-243"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

The CHORI-243 sheep (M) (Ovis aries) BAC library produced

by Pieter de Jong's lab at CHORI

<http://bacpac.chori.org/library.php?id=162>"

ORIGIN

Query Match 29.4%; Score 88.2; DB 13; Length 183;
Best Local Similarity 77.4%; Pred. No. 2.9e-12;
Matches 120; Conservative 0; Mismatches 33; Indels 2; Gaps 1;

QY 64 AACACGCTAACTCCCAATTTGCTACTACTAGGGAGTTTATCATGAAGAAATGTTAAA 123

DB 31 AACAGCGAGCTCCCAATTTGCTCATACTAGCGAAATCCATCAAGAAAATGTTGAAT 90

QY 124 TGCTTAATTTTCAATTAAGAAGTAAACACAGAGTGTGCAACATGCTGTTAAATAACCCGAC 183

DB 91 --CTATTTTAAATAAAGTACCAGAGTATGCAACATCTGTTAACTCAAAAC 148

QY 184 AACTTCAATCATATAGCTGTAGTAGTGCAAT 218

DB 149 AAATTCCTCTCATATAGCTGTAGTAGTGATTT 183

RESULT 3

BJ924615/c

LOCUS

DEFINITION BJ924615 MF01ALA cDNA Oryzias latipes cdna clone MF01ALA034p21 3',

mRNA sequence.

ACCESSION BJ924615

VERSION BJ924615.1

KEYWORDS GI:66764892

SOURCE EST.

ORGANISM Oryzias latipes (Japanese medaka)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

1 (bases 1 to 684)

Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.

Medaka EST Project in Takeda's lab

Unpublished (2001)

CONTACT: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tehini@genes.nig.ac.jp.

Location/Qualifiers

1. .684

/organism="Oryzias latipes"

/mol_type="mRNA"

/strain="Hd-r"

/db_xref="taxon:8090"

/clone="MF01ALA034p21"

/sex="mixture of female and male"

/tissue_type="adult liver"

/dev_stage="adult"

/clone_lib="MF01ALA CDNA"

ORIGIN

Query Match 15.3%; Score 45.8; DB 2; Length 684;

Best Local Similarity 53.0%; Pred. No. 0.36;

Matches 98; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 36 AAGACTTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTACCTAAGTAG 95

DB 325 AAGAACTAAATCCCTATATCTTCTGAATAATAAAAAGCTGTTAATATGTTATGATAG 266

QY 96 GCGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAATAAGAAGTAACAGAGT 155

DB 265 ATGAATAGACTTTTGAAAAAATGTTAAAAATGACAGAGTTGTGAAGTGTATTATCAAGAT 206

QY 156 GTTGCAACATGCTGTTAAATAAACCGACAACTTCAATCACTATAGCTGTAGTAGAGTGC 215

DB 205 TGAGGAAATACTCTCAAAAAACATTTCAACTTTTAATAATAAATTTGTTAATAAATTC 146

QY 216 ATTCT 220

DB 145 AATT 141

RESULT 4

BJ919954/c

LOCUS

DEFINITION BJ919954 MF01ALA cDNA Oryzias latipes cdna clone MF01ALA021h01 3',

mRNA sequence.

ACCESSION BJ919954

VERSION BJ919954.1

KEYWORDS GI:66760231

SOURCE EST.

ORGANISM Oryzias latipes (Japanese medaka)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

1 (bases 1 to 765)

Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.

Medaka EST Project in Takeda's lab

Unpublished (2001)

CONTACT: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tehini@genes.nig.ac.jp.

Location/Qualifiers

1. .765

/organism="Oryzias latipes"

/mol_type="mRNA"

/strain="Hd-r"

/db_xref="taxon:8090"

/clone="MF01ALA021h01"

/sex="mixture of female and male"

/tissue_type="adult liver"

One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 Plate: EK.2298 row: C column: 10
 High quality sequence stop: 513.
 Location/Qualifiers
 1..604
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="EK229834"
 /clone_lib="Exelixis FlyTag CK01 pcDNA-SK+"
 /note="Organ: mixed stage embryos, imaginal disks, and adult heads; Vector: pcDNA-SK+; Site 1: NotI; Site 2: XhoI; Random primed, normalized library from mixed stage embryos, imaginal disks, and adult heads."
 ORIGIN
 Query Match 14.5%; Score 43.6; DB 8; Length 604;
 Best Local Similarity 58.5%; Pred. No. 1.3;
 Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 29 TAGCATGAAGACTTGAATATCTTCATTTGAATAAACAAGCTAACTCCCAATTTGTACC 88
 Db 329 TATATTAACCTTAAAAATATTTTTTTTAAAAAAAATATTTATTATTAATTTACT 270
 QY 89 TAACTAGGGGAGTTTATCATGAAGAAATGTTAAATGCTTTAAATTTCAATAAGAAGTAA 148
 Db 269 TAATRAACTATTTTATATAATAATTTTATATAATAAAATTTATTAAATAATTATAA 210
 QY 149 CCAGAGTGTT 158
 Db 209 AAATATTTT 200
 RESULT 9
 AZ123198
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 Zhao, S., Nierman, W., Feidblyum, T., Malek, J., Shatsman, S.,
 Akinret, B., Levine, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
 Jong, P. and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Other GSSs: RPCI-23-22D9.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieter@jorg.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html
 Plate: 22 row: D column: 9
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..818

Genome Science Group (Genome Network Project Core Group).

The Transcriptional Landscape of the Mammalian Genome

Science 309, 1559-1563 (2005)

8 (bases 1 to 1933)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,

Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N.,

Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa, 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,

URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/.

Location/Qualifiers

1. .1933

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:A130072A19"

/db_xref="taxon:10090"

/clones="A130072A19"

/tissue_type="thymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days neonate"

1. .1933

/note="putative"

unclassifiable"

Query Match 14.0%; Score 42; DB 6; Length 1933;

Best Local Similarity 57.7%; Pred. No. 4;

Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 24 GTGACTACGAGAGCTTGAATATCTTCATTTGAATTAACAGCTCACTCCCAATT 83

|||||

Db 1214 GTGATTTTAAATATCTTAATGAATTTCTTAAGTTATAATAAAGCATGGTTTTTATT 1273

|||||

QY 84 GTACTAATAGGGGAGTTTATCATGAAGAATGTTTAAATGCTTAATTTCAAAATAGA 143

|||||

Db 1274 ATAGTTAAACATGTTAATTTTAAAGGAGTAATCTTAAAGGGCTAGTTTACTCTGAAGT 1333

|||||

QY 144 AGTAAACAGA 153

|||||

Db 1334 AGTAGTAAGA 1343

|||||

RESULT 12

BZ281627

LOCUS

DEFINITION

BZ281627 684 bp DNA linear GSS 15-OCT-2002

CH230-38403.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone

CH230-38403, genomic survey sequence.

ACCESSION

BZ281627.1 GI:24009141

VERSION

GSS.

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Rattus norvegicus

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 684)

Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,

Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,

Riggs, F., de Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: shao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 384 row: O column: 3

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .684

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SsNHsd/MCW"

/db_xref="taxon:10116"

/clones="CH230-38403"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 2"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match 13.8%; Score 41.4; DB 11; Length 684;

Best Local Similarity 48.1%; Pred. No. 5.1;

Matches 117; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 58 TGAATAAACAGCTAACTCCCAATTTGTACTACTAGGGAGTTTATCATGAAGAATG 117

|||||

Db 225 TGTAAATACAGATATTGTTGCAAAATTTATATGTAATTTGTCCATATGCTGCACAGAGATA 284

|||||

QY 118 TTTAAATGCTTAATTTTCAAAATAGAAGTAACCAAGAGTGTTCACCAATGCTGTTAAATAA 177

|||||

Db 285 AATATATCATTTAATGTTAT 344

|||||

QY 178 CCGGACAACTTCAATCACTATAGCTGTAGTAGTGCATTTCTGCAAGGATCCAGAGTA 237

|||||

Db 345 CAGATTTCTATAACCTATCTGATACCTGCTGCTATTATTTAAATTTAAAGAACTACTCTAATA 404

|||||

QY 238 ACCAGTATTTGGAAATGCAATGTTGAACCGACCATTAATTAATCTGCTGATTAGGAG 297

|||||

Db 405 ATTGAAGAATTTATAGATATTTATATATTTTACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 464

|||||

QY 298 TAT 300

|||||

Db 465 GAT 467

|||||

RESULT 13

DN248758

LOCUS

DEFINITION

DN248758 674 bp mRNA linear EST 29-JUN-2005

ACAE-aaa39e12.g1 Hydra EST UCI 5 Hydra magnipapillata cDNA 5', mRNA

sequence.

ACCESSION

DN248758


```

VERSION      DN248758.2  GI:68340052
KEYWORDS
SOURCE       Hydra magnipapillata
ORGANISM     Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
              Hydridae; Hydra.
REFERENCE    1 (bases 1 to 674)
AUTHORS      Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
              Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
              Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,
              Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
              Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
              WashU Hydra EST Project
              Unpublished (2002)
TITLE        On Mar 1, 2005 this sequence version replaced gi:60411218.
JOURNAL      Washington University School of Medicine
COMMENT      4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Library material provided by Hans Bode & Dirk Lindgens, Univ. of
              Calif., Irvine Library constructed by Dirk Lindgens Univ. of
              Calif., Irvine Library sequenced by Washington University Genome
              Sequencing Center For information on obtaining a clone please
              contact: Hans Bode (hrobode@uci.edu)
              original QR value of 876 was extended to value 877 (,)
              This trace has been recalled with phred
              original value before phred recall for SL was 91
              original value before phred recall for SR was 879
              Seq primer: -40RP from Gibco.
FEATURES     source
              Location/Qualifiers
                1..674
                  /organism="Hydra magnipapillata"
                  /mol_type="mRNA"
                  /strain="105"
                  /db_xref="taxon:6085"
                  /lab_host="DH10B"
                  /clone_lib="Hydra EST UCI 5"
                  /notes="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I;
                    a.1st strand cDNA was primed with a Not I primer-adaptor
                    (5' - PGACTAGTTCAGACGACGCCGCC(T)15-3')
                    b.Double-stranded cDNA was ligated to Sal I adaptor,
                    digested with Not I and cloned into the pSPORT 1-vector
                    pre-cut with Not I and Sal I. c.The ligation mix was
                    transformed into DH10B cells. d.The cells were grown in
                    SOC = 5% yeast, 20g tryptone, 0.5 g NaCl, 10 mM MgSO4, 10
                    mM MgCl, 0.2% glucose/liter, (no antibiotic). e.DMSO was
                    added to a final conc. of 10% as a cryoprotectant and
                    frozen f.The titre before freezing was determined as
                    ~2400/100 ul. Assuming a 10% loss upon freezing, the titre
                    is probably ~2100/ 100 ul. g.9 tubes each containing ~
                    2100 clones/100 ul (= total of ~19,000) are enclosed.
                    h.The frequency of vectors containing inserts is 96% as
                    determined by digestion check after picking 24 clones,
                    miniprep and subsequent digestion with Not I and Sal I.
                    i.A low level of 32p was used in the cDNA synthesis
                    procedure. The level measured by holding a Geiger Counter
                    next to a tube was background."
ORIGIN
Query Match      13.6%; Score 40.8; DB 9; Length 674;
Best Local Similarity 56.8%; Pred. No. 7.2;
Matches 75; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 56 TTGAAATAACAGCTAACTCCAAAATTGTACCTAACTAGGGAGTTTATCATGAGAAA 115
Db 418 TTTTAGCAACAGATTTCTTCCAAAGTTTCAAAATTAGCTATGAGTGATGATCATTAAGACA 477
QY 116 TGTTTAATGCTTAATTTTCAAAATAAGAACGATGATGTTGCAACATGCTGTTAAAT 175
Db 478 TTTTGTGATTTAAACATTTTCAATCTGGGGTTGCAAGACTGATAGTACTGATGTAAG 537

QY 176 AACCCGACAAAC 187
Db 538 CATTCATAAAC 549

RESULT 14
BH463846/c
LOCUS      BH463846      788 bp      DNA      linear      GSS 13-DEC-2001
DEFINITION BOGVW38TF BOGV Brassica oleracea genomic clone BOGVW38, genomic
              survey sequence.
ACCESSION  BH463846
VERSION    BH463845.1  GI:17658643
KEYWORDS   GSS.
SOURCE     Brassica oleracea
ORGANISM   Brassica oleracea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE  1 (bases 1 to 788)
AUTHORS    Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
              Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
              Whole genome shotgun sequencing of Brassica oleracea and its
              application to gene discovery and annotation in Arabidopsis
              Genome Res. 15 (4), 487-495 (2005)
JOURNAL    15805490
PUBMED
COMMENT     Other GSSs: BOGVW38TR
              Contact: Chris Town
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA.
              Tel: 301-838-3523
              Fax: 301-838-0208
              Email: cdtown@tigr.org
              DNA is from a doubled haploid provided by Tom Osborn.
              Seq primer: TF
              Class: sheared ends.
FEATURES     Location/Qualifiers
              1..788
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BOGVW38"
                /notes="BOGV"
                /notes="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
              genomic DNA inserted into PHOS1 using BstXI linkers"
ORIGIN
Query Match      13.5%; Score 40.4; DB 11; Length 788;
Best Local Similarity 47.3%; Pred. No. 9.4;
Matches 122; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 23 AGTGACTACGATGAAGACTTGAATAATTCTTCAATTTGAAATAAACAGCTAACTCCCAAT 82
Db 740 ACTGATTTCCGATTCAGATATATATTCATATATAGTTCARATTTGGTTGTTTGTTCGA 681
QY 83 TGTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAATTTTCAAAATAG 142
Db 680 AGTACTTAATTTTATTATTTTCTATAAAAAATTTGTAACAAAATATTATTATAATTT 621
QY 143 AAGTAACGAGAGTTTGCAACATGCTGTTAAATAACCCGACCAACTTCATCACTATAGC 202
Db 620 TAGTACAAATTCGTATTAAAAATTTATGCAATTTAAAGCTACCAATCGATATTTTACAAC 561
QY 203 TGTAGTAGAGTCGATTCCTGCAAGGATCCGAGATTAACAGTAACTTTTGGAAATGCAATGTT 262
Db 560 TGAATAATAATTTTTTTTTTATAAAATTTGTAATAGACCATGATTTTGTATGCGCTGAC 501
QY 263 GAACCCGACCATCTAATTT 280
Db 500 CTGCCAACGTTCTGTGAGT 483

```

```

RESULT 15
CWI44136/c
LOCUS
DEFINITION
104 536_11137783 148 34960.028 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11137783, genomic survey
sequence.
ACCESSION
CWI44136
VERSION
CWI44136.1 GI:54836683
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 694)
REFERENCE
AUTHORS
Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McMenamy, J., Smith, M., Holsman, H., Roe, B.A., Wiley, G., Korf, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddloh, J.A. and
Martensen, R.A.
TITLE
Sorghum genome sequencing by methylation filtration
JOURNAL
PLOS Biol. 3 (1), e13 (2005)
PUBMED
15660154
COMMENT
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 536 row: e column: 07
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 694.
Location/Qualifiers
FEATURES
source
1..694
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATX623"
/db_xref="taxon:4558"
/clone="11137783"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site: 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."
ORIGIN
Query Match 13.4%; Score 40.2; DB 13; Length 694;
Best Local Similarity 52.7%; Pred. No. 10;
Matches 87; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 13 CTAATGCTAGTGACTACGATGAGACTTGAATATTCTTCATTGAAATAACAGCTA 72
Db 529 CTCAGACCGGAATATTACTATCATATAATTAATAGTACTAGATTGTGTTGAGTTGTT 470
Qy 73 ACTCCCAATGTGCTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGTTAATT 132
Db 469 CAACCCGAGTTTGACCCCTGTTTGAGAGTTTATTATGCTCTAAAGATCAAGTACTTCTT 410
Qy 133 TTCAATAAAGAGTAACACGAGTGTTCACATCTGTGTAATAA 177
Db 409 TTCAATAAATAAAGAAATTCATCTTCGAGAGTATATGCAATTA 365
RESULT 16
CC272514/c
LOCUS
DEFINITION
CC272514 1172 bp DNA linear GSS 13-MAY-2003
CH261-91F1 RM1.1 CH261 Gallus gallus genomic clone CH261-91F1,
genomic survey sequence.
ACCESSION
CC272514
CWI44136
LOCUS
DEFINITION
104 536_11137783 148 34960.028 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11137783, genomic survey
sequence.
ACCESSION
CWI44136
VERSION
CWI44136.1 GI:54836683
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 694)
REFERENCE
AUTHORS
Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McMenamy, J., Smith, M., Holsman, H., Roe, B.A., Wiley, G., Korf, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddloh, J.A. and
Martensen, R.A.
TITLE
Sorghum genome sequencing by methylation filtration
JOURNAL
PLOS Biol. 3 (1), e13 (2005)
PUBMED
15660154
COMMENT
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 536 row: e column: 07
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 694.
Location/Qualifiers
FEATURES
source
1..694
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATX623"
/db_xref="taxon:4558"
/clone="11137783"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site: 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."
ORIGIN
Query Match 13.4%; Score 40.2; DB 13; Length 694;
Best Local Similarity 52.7%; Pred. No. 10;
Matches 87; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 13 CTAATGCTAGTGACTACGATGAGACTTGAATATTCTTCATTGAAATAACAGCTA 72
Db 529 CTCAGACCGGAATATTACTATCATATAATTAATAGTACTAGATTGTGTTGAGTTGTT 470
Qy 73 ACTCCCAATGTGCTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGTTAATT 132
Db 469 CAACCCGAGTTTGACCCCTGTTTGAGAGTTTATTATGCTCTAAAGATCAAGTACTTCTT 410
Qy 133 TTCAATAAAGAGTAACACGAGTGTTCACATCTGTGTAATAA 177
Db 409 TTCAATAAATAAAGAAATTCATCTTCGAGAGTATATGCAATTA 365
RESULT 17
CD288289/c
LOCUS
DEFINITION
CD288289 543 bp mRNA linear EST 27-MAY-2003
3.P14.abd POE14 (Day_14_pregnant_ovine_endometrium) Ovis aries
cDNA, mRNA sequence.
ACCESSION
CD288289
VERSION
CD288289.1 GI:31086332
KEYWORDS
EST.
SOURCE
Ovis aries (sheep)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Caprinae; Ovis.
1 (bases 1 to 543)
REFERENCE
AUTHORS
Gray, C.A., Abbey, C.A., Beremand, P.D., Choi, Y., Farmer, J.L.,
Adelson, D.L., Thomas, T.L., Bazer, F.W. and Spencer, T.E.
TITLE
Identification of Endometrial Genes Regulated by Early Pregnancy,
Progesterone, and Interferon Tau in the Ovine Uterus
JOURNAL
Biol. Reprod. 74 (2), 383-394 (2006)
PUBMED
16251498
CC272514.1 GI:30627693
GSS.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1172)
REFERENCE
AUTHORS
Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
JOURNAL
Unpublished (2003)
COMMENT
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TAGCACTCATATAGGAGAGA
Class: BAC ends
High quality sequence start: 115
High quality sequence stop: 316.
Location/Qualifiers
FEATURES
source
1..1172
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-91F1"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTACBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"
ORIGIN
Query Match 13.4%; Score 40.2; DB 12; Length 1172;
Best Local Similarity 54.4%; Pred. No. 11;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
Qy 90 AACTAGGGGATTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATGAAGTAAC 149
Db 950 AAAAATGGGTGAACAAACAAAGACATGTTTCTTATTAATTTGCAAGTGACACTAGAA 891
Qy 150 CAGAGTGTGCAACATCTGTTAATAACCCGACAACTTCAATCACTATAGCTGTAGTA 209
Db 890 AAGATTTTGAACATCGGTTTACAAATCGGCCATTTTTTTTCCCAAGGAAAAA 831
Qy 210 GAGTGCAATCTTCAAGGATCCAGAGTAA 238
Db 830 AACGAATTAACAGTAGTAGAAGTAA 802
RESULT 17
CD288289/c
LOCUS
DEFINITION
CD288289 543 bp mRNA linear EST 27-MAY-2003
3.P14.abd POE14 (Day_14_pregnant_ovine_endometrium) Ovis aries
cDNA, mRNA sequence.
ACCESSION
CD288289
VERSION
CD288289.1 GI:31086332
KEYWORDS
EST.
SOURCE
Ovis aries (sheep)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Caprinae; Ovis.
1 (bases 1 to 543)
REFERENCE
AUTHORS
Gray, C.A., Abbey, C.A., Beremand, P.D., Choi, Y., Farmer, J.L.,
Adelson, D.L., Thomas, T.L., Bazer, F.W. and Spencer, T.E.
TITLE
Identification of Endometrial Genes Regulated by Early Pregnancy,
Progesterone, and Interferon Tau in the Ovine Uterus
JOURNAL
Biol. Reprod. 74 (2), 383-394 (2006)
PUBMED
16251498

```

```

COMMENT      Contact: Thomas E. Spencer
              Center for Animal Biotechnology and Genomics
              Texas A&M University
              Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
              USA
              Tel: 9798454896
              Fax: 9798622662
              Email: tspender@ansc.tamu.edu.
              Location/Qualifiers
FEATURES     source
              1..543
              /organism="Ovis aries"
              /mol_type="mRNA"
              /db_xref="taxon:9940"
              /sex="Female"
              /tissue_type="endometrium"
              /dev_stage="Day 14 pregnant"
              /clone_lib="POE14 (Day 14 pregnant ovine endometrium)"
              /notes="Organ: uterus; Vector: Triplex2; Site: 1: EcoRI;
              Site2: XhoI; Non-normalized library, sequenced 5' with
              Triplex2 primer (CTCCGAGATCTGACGAGC). Library constructed
              by Clontech with total RNA extracted using the Trizol
              method and pooled from 5 females."
ORIGIN
Query Match      13.2%; Score 39.6; DB 5; Length 543;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 15 AAATGCTGAAGTACTACGATGAAGACCTTGAATAATCTTTCATTTGAATAAAGACGCTAAC 74
DB 389 ATAGTCTTTGGCTACATGAATTAATAATTAGAATCAGCAATAGAAACATCTGGGAAA 330
QY 75 TCCCAATTTGACTACTAGGAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTT 134
DB 329 TTCCCAATATCAGGAATTTAAACAAAGCAGCTTTTAAATATATATATTAATTAATCTAT 270
QY 135 CAATAAGAGTAGTACGAGAGTGTGCAACATGCTGTTAAATATACCCGACAACTTCAATC 194
DB 269 CAAAGAGACATCAAGAAGAAATGAGAAAATATTTTAAAGTAAATGCAAAAAATTTCTTC 210
QY 195 ACTATAGCTCTAGTAGAG 212
DB 209 AATATATCGAATTATAG 192

RESULT 18
LOCUS      CJ361302 712 bp mRNA linear EST 15-JUN-2005
DEFINITION      embryo Molgula tectiformis unpublished cDNA library, cleaving
VERSION      CJ361302.1 GI:67801450
KEYWORDS      EST.
SOURCE      Molgula tectiformis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
              Molgula tectiformis
              1 (bases 1 to 712)
              Gyoja, F., Satou, Y. and Satoh, N.
              Expressed genes in Molgula tectiformis
              Unpublished (2005)
              Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp
              When you want to obtain this EST clone, please send an e-mail to
              Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka
              Satou (yutaka@ascidian.zool.kyoto-u.ac.jp).
              Location/Qualifiers
FEATURES     source
              1..712

Contact: Thomas E. Spencer
Center for Animal Biotechnology and Genomics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798454896
Fax: 9798622662
Email: tspender@ansc.tamu.edu.
Location/Qualifiers
FEATURES     source
              1..543
              /organism="Ovis aries"
              /mol_type="mRNA"
              /db_xref="taxon:9940"
              /sex="Female"
              /tissue_type="endometrium"
              /dev_stage="Day 14 pregnant"
              /clone_lib="POE14 (Day 14 pregnant ovine endometrium)"
              /notes="Organ: uterus; Vector: Triplex2; Site: 1: EcoRI;
              Site2: XhoI; Non-normalized library, sequenced 5' with
              Triplex2 primer (CTCCGAGATCTGACGAGC). Library constructed
              by Clontech with total RNA extracted using the Trizol
              method and pooled from 5 females."
ORIGIN
Query Match      13.2%; Score 39.6; DB 5; Length 543;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 15 AAATGCTGAAGTACTACGATGAAGACCTTGAATAATCTTTCATTTGAATAAAGACGCTAAC 74
DB 389 ATAGTCTTTGGCTACATGAATTAATAATTAGAATCAGCAATAGAAACATCTGGGAAA 330
QY 75 TCCCAATTTGACTACTAGGAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTT 134
DB 329 TTCCCAATATCAGGAATTTAAACAAAGCAGCTTTTAAATATATATATTAATTAATCTAT 270
QY 135 CAATAAGAGTAGTACGAGAGTGTGCAACATGCTGTTAAATATACCCGACAACTTCAATC 194
DB 269 CAAAGAGACATCAAGAAGAAATGAGAAAATATTTTAAAGTAAATGCAAAAAATTTCTTC 210
QY 195 ACTATAGCTCTAGTAGAG 212
DB 209 AATATATCGAATTATAG 192

RESULT 19
LOCUS      CT404992 743 bp DNA linear GSS 03-NOV-2005
DEFINITION      Sus scrofa genomic clone CH242-289D18, genomic survey sequence.
VERSION      CT404992.1 GI:80063530
KEYWORDS      GSS.
SOURCE      Sus scrofa (pig)
ORGANISM      Sus scrofa
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
              Sus.
              1 (bases 1 to 743)
              Humphray, S.J., Plumb, R.W. and Durham, J.L.
              Direct Submission
              Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              humquery@sanger.ac.uk Unpublished
              This sequence was generated from the SP6 end of BAC 289D18. 289D18
              is part of the CHORI-242 BAC Library created by P. de Jong. Further
              details: http://www.sanger.ac.uk/Projects/S_scrofa/.
              Location/Qualifiers
FEATURES     source
              1..743
              /organism="Sus scrofa"
              /mol_type="genomic DNA"
              /db_xref="taxon:9823"
              /clone="CH242-289D18"
              /tissue_type="White blood cells"
              /note="vector pTABAC1.3_BamHI
              sex female"
ORIGIN
Query Match      13.2%; Score 39.6; DB 14; Length 743;
Best Local Similarity 58.5%; Pred. No. 15;
Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 41 TTGAAATATTTCTTCATTTGAAATAAAGACAGCTAACCTCCAAATTTGTACCTAATAGGGAG 100
DB 24 TTGAAGTATAGTTGATTTACAATATTCAGTTAGTTTCAAGTGATATAGCAATGTGTTTCAAG 83
QY 101 TTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGT 158
DB 84 TTATTCATAATATATATATAATAATATATATATATATATATATATATATATATATATATAT 141

RESULT 20

```

CJ381709
LOCUS CJ381709 745 bp mRNA linear EST 15-JUN-2005
DEFINITION CJ381709 Molgula tectiformis unpublished cDNA library, mixture of gastrula and neurula Molgula tectiformis cDNA clone mtga015g03 3', mRNA sequence.
ACCESSION CJ381709 GI:67822106
VERSION
KEYWORDS
SOURCE
ORGANISM Molgula tectiformis
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
TITLE 1 (bases 1 to 745)
JOURNAL Gyoja, F., Satou, Y. and Satoh, N.
COMMENT Expressed genes in Molgula tectiformis
Unpublished (2005)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
When you want to obtain this EST clone, please send an e-mail to
Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka
Satou (yutaka@ascidian.zool.kyoto-u.ac.jp).

FEATURES
source
1. 745
/organism="Molgula tectiformis"
/mol_type="mRNA"
/db_xref="taxon:30286"
/clone="mtga015g03"
/tissue_type="whole animal"
/dev_stage="mixture of gastrula and neurula"
/clone_lib="Molgula tectiformis unpublished cDNA library,
mixture of gastrula and neurula"

ORIGIN
Query Match 13.2%; Score 39.6; DB 5; Length 745;
Best Local Similarity 53.2%; Pred. No. 15;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 4 ATTGACAGGCTTAATGCTAGTACTAGTACGAGCTTGAAGACTTGAATATTTCTTCATTTGAAAT 63
DB 80 AATGAAATGAATGCTCAGTTTATGGAAGCAACATTTATTTTAACTAACTTTAAAAAC 139
QY 64 AAACAGCTACTCCCAATTTGTACTTAACCTAGGAGTGTATCATGAGAAATGTTTAAA 123
DB 140 AACTGTGAACCTCATTAATTTTGACATTTGTTGTTTATGTTATGTTGTTAAAGTAAAC 199
QY 124 TGCTTAATTTTCAAATGAAGTAACCAAGAGTGTGCA 161
DB 200 TGCTTAAGTTTCAGAAACAGTTAATGGCAGTAGTGTA 237

RESULT 21
CJ385948
LOCUS CJ385948 810 bp mRNA linear EST 15-JUN-2005
DEFINITION CJ385948 Molgula tectiformis unpublished cDNA library, mixture of gastrula and neurula Molgula tectiformis cDNA clone mtga030j02 3', mRNA sequence.
ACCESSION CJ385948
VERSION
KEYWORDS
SOURCE
ORGANISM Molgula tectiformis
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
TITLE 1 (bases 1 to 810)
JOURNAL Gyoja, F., Satou, Y. and Satoh, N.
COMMENT Expressed genes in Molgula tectiformis
Unpublished (2005)
Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
When you want to obtain this EST clone, please send an e-mail to
Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka
Satou (yutaka@ascidian.zool.kyoto-u.ac.jp).

FEATURES
source
1. 810
/organism="Molgula tectiformis"
/mol_type="mRNA"
/db_xref="taxon:30286"
/clone="mtga030j02"
/tissue_type="whole animal"
/dev_stage="mixture of gastrula and neurula"
/clone_lib="Molgula tectiformis unpublished cDNA library,
mixture of gastrula and neurula"

ORIGIN
Query Match 13.2%; Score 39.6; DB 5; Length 810;
Best Local Similarity 53.2%; Pred. No. 15;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 4 ATTGACAGGCTTAATGCTAGTACTAGTACGAGCTTGAAGACTTGAATATTTCTTCATTTGAAAT 63
DB 77 AATGAAATGAATGCTCAGTTTATGGAAGCAACATTTATTTTAACTAACTTTAAAAAC 136
QY 64 AAACAGCTACTCCCAATTTGTACTTAACCTAGGAGTGTATCATGAGAAATGTTTAAA 123
DB 137 AACTGTGAACCTCATTAATTTTGACATTTGTTGTTTATGTTATGTTGTTAAAGTAAAC 196
QY 124 TGCTTAATTTTCAAATGAAGTAACCAAGAGTGTGCA 161
DB 197 TGCTTAAGTTTCAGAAACAGTTAATGGCAGTAGTGTA 234

RESULT 22
CJ2012938
LOCUS CJ2012938 539 bp DNA linear GSS 10-JAN-2005
DEFINITION CJ2012938 CH240_507D02.TJ CHORI-240 Bos taurus genomic clone CH240_507D02,
genomic survey sequence.
ACCESSION CJ2012938
VERSION CJ2012938.1 GI:57341499
KEYWORDS GSS.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
TITLE Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 539)
Costa, J.N., Mota, M. and Caetano, A.R.
Brazil's Contribution to End-Sequencing the Bovine BAC Library
CHORI-240
Unpublished (2003)
Other GSSs: CH240_507D02.TV
Contact: Caetano AR
Department of Biotechnology
Embrapa Recursos Geneticos e Biotecnologia
Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P.
02372, 70770-900 Brasil
Tel: 55 61 448 4778
Fax: 55 61 340 3658
Email: acaetano@cenargen.embrapa.br
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). Bases shown have Phred
quality value equal to or higher than 20. Bases with quality value
below 20 were masked with 'N'. For BAC library availability, please
contact Pieter de Jong (pdejong@mail.cho.org). Clones may be
purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping

ampicillin."

ORIGIN
Query Match 12.9%; Score 38.8; DB 12; Length 547;
Best Local Similarity 50.5%; Pred. No. 23;
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 4 ATTGACAGCTAAATGCTAAGTACTAGTACGATGAGACTTGAAATATCTTCATTTGAAAT 63
DB |||||
363 ATACATATGATATATTTATTTATTTCAATATATAAACACCTGAGCATGATGTTATCTCGTT 304
QY 64 AAACAGCTAACTCCCAATTTGACCTAAGTACTAGGAGTTCATGAGAGAAATGTTTAAA 123
DB |||||
303 AGACATCTACGCCGCTTTTAAACATTTTACATGAGTTAAATTTTAAAGAAATGTAAGGA 244
QY 124 TGCTTAAATTTCAATAAGAGTACACAGAGTGTTCACACATGCTGTTTAAATAACCCGAC 183
DB |||||
243 TTTTITTTTGAAGAGATGATGAGGACGCTGTTGAACTAGTCTTTTAACTACAGTAG 184
QY 184 AACTT 189
DB |||||
183 AGATTT 178

RESULT 25
BH929322/c
LOCUS
DEFINITION BH929322.1 B.oleracea002 Brassica oleracea genomic, genomic survey
ACCESSION BH929322
VERSION BH929322.1 GI:23409388
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 690)
AUTHORS Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
Nash, W., Rabinowicz, P.D. and Wilson, R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: odi02 row: d column: 08
Seq primer: -21UPPOT forward
Class: shotgun
High quality sequence start: 55
High quality sequence stop: 543.
FEATURES
source
1..690
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea 101000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match 12.9%; Score 38.8; DB 11; Length 690;
Best Local Similarity 50.5%; Pred. No. 24;
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 38 GACTTGAATATTTCTTCATTTTGAATAAACAGCTAACCTCCCAATTTGACTACTAGGG 97
DB |||||
224 GATTGTGTTATTTTAAAAAATAACAAATTAACCTAACCTGTTTAAAAATAATA 165

QY 98 GAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAAATAAGAACTAACACAGTGT 157
DB |||||
164 TAATAAATAATACCAATAAATTTAGATTTTACAAATTTAAATTTATCAAAATCTGTGATT 105
QY 158 TGCAACATGCTGTTTAAATAACCCGACAACTTCATCTACTATAGCTGTAGTAGATGAT 217
DB |||||
104 ATTATTTTAAATAAATAACAATAACATATTTTAAATTTATAGATTTTATAAAGAGATT 45
QY 218 TCTGCA 223
DB |||||
44 CCACCA 39

RESULT 26
BF336378
LOCUS
DEFINITION BF336378.1 Homo sapiens linear
ACCESSION BF336378
VERSION BF336378.1 GI:11307126
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-CT0510-
010900-319-c09&t3=2000-09-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 3
High quality sequence stop: 359.
FEATURES
source
1..359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0510"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN

Query Match 12.9%; Score 38.6; DB 7; Length 359;
Best Local Similarity 56.5%; Pred. No. 25;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 41 TTGAATATTTCTTCATTTTGAATAAACAGCTAACCTCCCAATTTGACTACTAGGGAG 100

```

RESULT 28
CR068462/c
LOCUS
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP272k04, genomic survey sequence.
ACCESSION CR068462
VERSION 1
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 535)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
FEATURES
source Location/Qualifiers
1..535
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP272k04"
/clone_lib="MHPP"
ORIGIN
Query Match 12.8%; Score 38.4; DB 14; Length 535;
Best Local Similarity 58.9%; Pred. No. 30;
Matches 66; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Qy 25 TGACTGACGTGAAGACTTGAATAATCTTTCATTTGAATAAACAAGCTAACTCCCAATTG 84
Db 160 TGAATTCAAATCGCGCTTAAATATGTCAGTGGACAGAAAAAATTCATACAGATTA 101
Qy 85 TACCTAACTAGGGAGTTCATCATGAAGAAATGTTAAATGCTTAATTTCA 136
Db 100 TATTAACTAGGGCAATGTTGGTACAGATTTTTTAAAGGCATAATGTCA 49
RESULT 29
CL341203/c
LOCUS
DEFINITION RPCI44_260K19.r RPCI-44 Sus scrofa genomic clone RPCI44_260K19,
genomic survey sequence.
ACCESSION CL341203
VERSION 1
KEYWORDS GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE 1 (bases 1 to 683)
AUTHORS Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,
Beever,J.E. and Schook,L.B.
TITLE Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
Through Comparative Genomics
JOURNAL Unpublished (2004)
COMMENT Other GSSs: RPCI44_260K19.f
Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu
Clones are derived from the porcine BAC library RPCI-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library

```


[illegible]

ACCESSION

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 CA973570.1 GI:58787092
 xenopus tropicalis (western clawed frog)
 xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 779)
 Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute
 Unpublished (2004)
 Other_ESTs: JGI_CAAP8099.fwd
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Robert M. Grainger
 cDNA Library Preparation: Bruce Blumberg Laboratory, University of
 California, Irvine
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>

clone Distribution: I.M.A.G.E. Consortium/LLNL
 http://image.llnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone id and the direction of sequencing. The suffix '.rev'
 indicates a reverse sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Poly-A: Based upon the presence of a run of 14 or more T residues
 at the beginning of the sequence, this clone was polyadenylated.
 The resulting Poly-T sequence has been removed.
 Plate: CAAP 0085 row: e column: 9
 High quality sequence stop: 779
 POLYA=Yes.

Location/Qualifiers
 1..779
 /organism="Xenopus tropicalis"

FEATURES
 source

```

organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7714547"
/tissue_type="Intestine"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropint1"
/notes="vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT

```


restriction digestion to determine average insert size (96 clones) and by sequencing (-4 96-well plates) to confirm library quality (e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core Facility at: bovine@net.missouri.edu. Bonaldo MF, Lennon G, Soares MB. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res. 1996; 6:791-806. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Bonaldo, P Jelene, L Su, L Lawton, A Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG TISSUE=Day 14 CL from a pregnant animal TAG_SEQ=CAGGTGAGAT"

ORIGIN

Query Match 12.7%; Score 38.2; DB 8; Length 756;
Best Local Similarity 57.6%; Pred. No. 35;
Matches 87; Conservative 0; Mismatches 63; Indels 1; Gaps 1;
Qy 15 AAATGCTAAGTACGATGACGACTTGAATATCTTCTTCAATTTGAATTAACAGCTAAC 74
Db |||||
Qy 535 AAATGCTAAGTACGATTAACATGATGATGATATTTTATCTGACATCAGATGATTT 476
Db |||||
Qy 75 TCCCAATTTGACCT-AACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
Db |||||
Qy 475 TAACGTATATTAACTGAAAGAAAGAAATTTAAGCAGATGATATAATATCTTATTTT 416
Db |||||
Qy 134 TCAATAAGAAGTAAACAGAGTGTTCGAC 164
Db |||||
Qy 415 TATTAAATAACTAAAGAAAGTGTTCCTGCA 385
Db |||||

RESULT 33

BH508774
LOCUS BH508774 759 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGQ Brassica oleracea genomic clone BOGQ34, genomic survey sequence.
ACCESSION BH508774
VERSION BH508774.1 GI:17716864
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 759)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other GSSs: BOGQ34TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.

FEATURES

source
1..759
Location/Qualifiers
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOGQ34"
/clone_lib="BOGQ"
/note="vector: PHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 12.7%; Score 38.2; DB 11; Length 759;
Best Local Similarity 48.4%; Pred. No. 35;
Matches 106; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
Qy 13 CTAATGCTAAGTACGATGACGACTTGAATATCTTCAATTTGAATTAACAGCTA 72
Db |||||
Qy 29 CCACATTTCTACATTTGATTAATGATGATTTAACTCTCTTCATGGTATATGATCGT 88
Db |||||
Qy 73 ACTCCCAATTTGACCTAACTAGGGGATTTATCATGAAGAAATGTTTAAATGCTTAAT 132
Db |||||
Qy 89 TATAATTTTTTATAAATTAATTTAGTAAACCTTCATAATACTCCATTAATCGGTGAAT 148
Db |||||
Qy 133 TTCAATAAGAAGTAAACGAGAGTGTTCACATGCTTTAAATACCCGACAACTTCAA 192
Db |||||
Qy 149 AACGACTATAAGTCACAAATTTCTTTGAAAACCGAGATACAAAGGCTCCAAACCTCT 208
Db |||||
Qy 193 TCACATAGCTGTAGTAGAGTGCAATTTCTGCAAGGATGCC 231
Db |||||
Qy 209 TCAACATCATCATATGTTAGTGCATGATGAACATATGCC 247
Db |||||

RESULT 34

CZ764379/c
LOCUS CZ764379 808 bp DNA linear GSS 26-JUL-2005
DEFINITION OC_Ba0119E01.f OC_Ba Oryza coarctata genomic clone OC_Ba0119E01 5', genomic survey sequence.
ACCESSION CZ764379
VERSION CZ764379.1 GI:71204230
KEYWORDS GSS.
SOURCE Oryza coarctata (Porteresia coarctata)
ORGANISM Oryza coarctata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 808)
AUTHORS Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and Wing,R.
TITLE OMAP (Oryza Map Alignment Project) - Arizona Genomics Institute
JOURNAL Unpublished (2005)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0119 row: E column: 01
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1..808
/organism="Oryza coarctata"
/mol_type="genomic DNA"

```

/db_xref="taxon:77588"
/clone="OC_Ba0119E01"
/tissue_type="leaves"
/dev stage="mature"
/lab_host="DH10B"
/clone_lib="OC_Ba"
/notes="vector: PAGIBAC1, Site_1: HindIII, Site_2: HindIII"

```

ORIGIN

```

Query Match      12.7%; Score 38.2; DB 13; Length 808;
Best Local Similarity 52.1%; Pred. No. 35;
Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 66 ACAGCTAACTCCCAATGTGACTTAAGTGGGGAGTTTATCATGAAGAAATGTTTAAATG 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 640 ACATCTCTCCACCAACAAAGGACATGATCTGAATTTATGAGCAACACTGTTGAATAA 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 126 CTTAATTTTCAATAAGTAAGTACACAGAGTGTTCACATGCTGTTAAATAACCCGACAA 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 580 TAAAGTTGTCCAAATGAGATTATACCCAGGGGATAGAAACACACTGTTTATAACACTTACAT 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 186 ACTTCATCACTACTAGCTAGTAGTGCATTCCTGCAAGGAT 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 520 AATGCTCAATATATGCGCTTTATTTGGATCTCAATCAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 35

```

DU208665
LOCUS      DU208665      818 bp      DNA      linear      GSS 04-OCT-2005
DEFINITION 198574240869 CHORI-243 Ovis aries genomic clone CH243-402L23,
genomic survey sequence.

```

ACCESSION

```

DU208665
VERSION    DU208665.1 GI:76915605

```

KEYWORDS

```

GSS.

```

SOURCE

```

Ovis aries (sheep)

```

ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Caprinae; Ovis.

```

REFERENCE

```

1 (bases 1 to 818)

```

AUTHORS

```

Cokkett, N., Shetty, J., de Jong, P., McEwan, J.C., Oddy, H. and

```

TITLE

```

Ovine BAC End Sequences from Library CHORI-243

```

JOURNAL

```

Unpublished (2004)

```

COMMENT

```

Contact: Ewen Kirkness
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-795-7536
Email: ekirknes@tigr.org
Sequences generated at the J. Craig Venter Institute Joint
Technology Center (JCVI/JTC; http://www.venterlinstitute.org/).
Original Trace: 198574240869 Trace.TI: gnl|ti|918735517
Insert Length: 184000 Std Error: 0.00 row: L column: 23
Seq primer: T7
Class: BAC ends.

```

FEATURES

```

source

```

```

1..818
    Location/Qualifiers
      /organism="Ovis aries"
      /mol_type="genomic DNA"
      /strain="Texel breed"
      /db_xref="taxon:9940"
      /clone="CH243-402L23"
      /sex="Wale"
      /cell_type="Blood"
      /clone_lib="CHORI-243"

```

```

/notes="vector: PRARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
The CHORI-243 sheep (M) (Ovis aries) BAC library produced
by Pieter de Jong's lab at CHORI
http://bacpac.chori.org/library.php?id=162"

```

ORIGIN

```

Query Match      12.7%; Score 38.2; DB 13; Length 818;

```

```

Best Local Similarity 56.9%; Pred. No. 35;
Matches 70; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

```

QY 41 TTGAATATTTCTTCATTGAAATATAACAGCTAACTCCCAATTTGTACTAACTAGGGAG 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 679 TTGGAATTCCTTCATCTTTAAATAGTTAGCTAACAGTTTTAAAGTCCACTCGAGGAG 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 TTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATGAAGTAACAGAGTGTTC 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 739 TTTACCTTAAGGGAAGTTTGTGTAATGCTCTGTTAAAAATGAAGTAACCTTTTCAGTAAC 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 AAC 163
    |||
DB 799 AAC 801
    |||

```

RESULT 36

```

AF156725
LOCUS      AF156725      3792 bp      DNA      linear      GSS 06-NOV-2000
DEFINITION Human Homo sapiens genomic, genomic survey sequence.

```

ACCESSION

```

AF156725
VERSION    AF156725.1 GI:5731122

```

KEYWORDS

```

GSS.

```

SOURCE

```

Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

REFERENCE

```

1 (bases 1 to 3792)

```

AUTHORS

```

Bepko, G., O'Brian, K.C., Kim, Y.C., Schreiber, G. and Pitterle, D.M.
A 1.4-Mb high-resolution physical map and contig of chromosome
segment 11p15.5 and genes in the L0H1A metastasis suppressor
region

```

JOURNAL

```

Genomics 55 (2), 164-175 (1999)

```

PUBMED

```

9933563

```

COMMENT

```

Contact: Bepko, G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
Kim, Y.C., Cao, Y., Pitterle, D.M., O'Brian, K.C. and Bepko, G.
SSA/Ro52 gene and expressed sequence tags in an 85 kb region of
chromosome segment 11p15.5
EcoRI fragment which connects clone pTWB9n12.03 (AF101866) with
pTWB113.08 (AF101895)
Class: unknown.

```

FEATURES

```

source

```

```

1..3792
    Location/Qualifiers
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /map="11p15.5; between beta-hemoglobin gene cluster and
the RRM1 gene"
      /clone_lib="Human"

```

ORIGIN

```

Query Match      12.7%; Score 38.2; DB 11; Length 3792;
Best Local Similarity 48.0%; Pred. No. 43;
Matches 109; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 39 ACTTGAATATTTCTTCATTGAAATAAACAGTAACTCCCAATTTGTACCTAACTAGGGG 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2629 AATTCAGCCATTCACATTTGATATGATATCAAGAAATCATATGTATATAAACTACTTG 2688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 99 AGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATGAAGTAACAGAGTGT 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2689 TCAATTTTAAATAATTAATAGAAAAGTATCAGCATAACTACCTAGTATAGGACGTT 2748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 159 GCAACATGCTGTTAAATAACCCGACAACTTCATCACTATAGCTGTAGTAGTGCAATT 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2749 ACAACATGCTCTATGAATGAGCTAAAGATGTGAGATGATCTTGTAGATCTTAAGTACATC 2808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 219 CTGCAAGGATCCAGAGTAACCACTATTTTGGAAATGCAATGTTCAA 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 2809 ACCATCAGTCAAGAAGGAGGATTCCTGCGCAAAAAATCAAGAA 2855

RESULT 37
A0836871
LOCUS
DEFINITION
HS_5463_A1_F01_T7A_RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plates=1039 Col=1 Row=K, genomic survey sequence.
ACCESSION
A0836871
VERSION
A0836871.1 GI:5806745
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 559)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1039 row: K column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 559.
FEATURES
Location/Qualifiers
source
1..559
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1039 Col=1 Row=K"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"
ORIGIN
Query Match 12.7%; Score 38; DB 11; Length 559;
Best Local Similarity 55.2%; Pred. No. 38;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 16 AATGCTAGTGACTACGATGAAGACTTGAAATATCTTCATTGCAATAAACAGCTAACT 75
|||||
Db 205 AATGGTGAGGATGTCGCTGATGTTCTATTATTAATGATCTGTATAGACAGTTAAGA 264
|||||

Qy 76 CCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAATGCTTAATTTTC 135
|||||
Db 265 TACATATGGTGTCTGATGTGTGTGTGTTATTCATTAATGAAGTGTGATTTATTATTTTC 324
|||||

Qy 136 AATAAAGAAGTAAC 149
|||||
Db 325 TACTATTAAATTAC 338
|||||

```

```

RESULT 38
DW602832/c
LOCUS
DEFINITION
CGX162-D05.xld-t SHGC-CGX2 Gasterosteus aculeatus cDNA clone
CGX162-D05 3', mRNA sequence.
ACCESSION
DW602832
VERSION
DW602832.1 GI:85192841
KEYWORDS
EST.
SOURCE
Gasterosteus aculeatus (three spined stickleback)
ORGANISM
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1091)
Kingsley,D.M., Peichel,C., Knecht,A., Balabhadra,S., Grimwood,J.,
Dickson,M., Schmutz,J., and Myers,R.M.
Expressed sequence tags from Gasterosteus aculeatus (2004)
Unpublished (2004)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 162
High quality sequence start: 30
High quality sequence stop: 901.
High quality sequence stop: 901.
FEATURES
Location/Qualifiers
source
1..1091
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CGX162-D05"
/sex="mixed male and female"
/tissue_type="eyes"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CGX2"
/note="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTTCGATCGGAGCGGCCCTT)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna library construction fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria."
ORIGIN
Query Match 12.7%; Score 38; DB 10; Length 1091;
Best Local Similarity 55.5%; Pred. No. 41;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 19 GCTAAGTGACTACGATGAAGACTTGAATATCTTCATTGGAATAACAGCTAACTCCC 78
|||||
Db 1067 GGTAAAGCGAACCTATGGAACCTGGTGTGTTGCTATAGTGACATGACCACTTTTACTGCA 1008
|||||

Qy 79 AAATTGTACTTACTAGGGGAGTTTATCATGAAGAAATGTTTAATGCTTAATTTTCAA 138
|||||
Db 1007 GAAGNGAGCATACCTCTTGTCTTTTACCATGCCTCAATGCATTTTATGCTTAGTCTCAA 948
|||||

```

```

QY      139 TAAGAAGT 146
Db      947 GAAGATGT 940
|||||

RESULT 39
CZ953670      914 bp DNA linear GSS 11-AUG-2005
DEFINITION   265847 Tomato EcoRI BAC Library Lycopersicon esculentum genomic
              clone SL_EcoRI0057L14 5, genomic survey sequence.
ACCESSION   CZ953670
VERSION     CZ953670.1 GI:72299235
KEYWORDS    GSS.
ORGANISM    Lycopersicon esculentum (Solanum lycopersicum)
            Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 914)
AUTHORS     Mueller,L.A., Buels,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
            Van Eck,J. and Stack,S.
TITLE       BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL     Unpublished (2005)
COMMENT     Other_GSSs: 273838
            Contact: Lukas Mueller
            Tanksley Lab, Dept. of Plant Breeding
            Cornell University
            251 Emerson Hall, Ithaca, NY 14853, USA
            Tel: 607-255-6557
            Fax: 607-255-6683
            Email: sgn-feedback@sgn.cornell.edu
            Plate: 57 row: L column: 14
            Seq primer: T7
            Class: BAC ends
            High quality sequence start: 6
            High quality sequence stop: 818.
            Location/Qualifiers
                source          1..914
                    /organism="Lycopersicon esculentum"
                    /mol_type="genomic DNA"
                    /cultivar="Heinz 1706"
                    /db_xref="taxon:4081"
                    /clone="SL_EcoRI0057L14"
                    /lab_hosts="E. coli"
                    /clone_lib="Tomato EcoRI BAC Library"
                    /note="Vector: unk; Site_1: EcoRI"

FEATURES             source
Query Match          12.6%; Score 37.8; DB 13; Length 914;
Best Local Similarity 54.7%; Pred. No. 45;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      10 AGGCTAAATGCTAGTACGATGAGACTTCGAAATATTTCTTCATTGAAATAAACAG 69
Db      554 ATGGAATAAACTCAATAAAGCGCTCAAACTCAAACTCAATCATATGAGATAAAAG 613
|||||

QY      70 CTAATCTCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTA 129
Db      614 TGACCTTCCAAAGCTTACCACACTAATGAAATTTAATTCGCGATAGTTACAAAAATA 673
|||||

QY      130 ATTTTCAAAATAAGAAGT 146
Db      674 TTCACCAAAATCCAAATT 690
|||||

RESULT 40
DU931319      950 bp DNA linear GSS 03-JAN-2006
DEFINITION   353827 Tomato EcoRI BAC Library Lycopersicon esculentum genomic
              clone SL_EcoRI0062K06 3, genomic survey sequence.
ACCESSION   DU931319
VERSION     DU931319.1 GI:84252534
KEYWORDS    GSS.

```

```

SOURCE          Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM        Lycopersicon esculentum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE       1 (bases 1 to 950)
AUTHORS         Mueller,L.A., Buels,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
                Van Eck,J. and Stack,S.
TITLE           BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL         Unpublished (2005)
COMMENT         Other_GSSs: 269087
                Contact: Lukas Mueller
                Tanksley Lab, Dept. of Plant Breeding
                Cornell University
                251 Emerson Hall, Ithaca, NY 14853, USA
                Tel: 607-255-6557
                Fax: 607-255-6683
                Email: sgn-feedback@sgn.cornell.edu
                Plate: 62 row: K column: 6
                Seq primer: SP6
                Class: BAC ends
                High quality sequence start: 54
                High quality sequence stop: 564.
                Location/Qualifiers
                    source          1..950
                        /organism="Lycopersicon esculentum"
                        /mol_type="genomic DNA"
                        /cultivar="Heinz 1706"
                        /db_xref="taxon:4081"
                        /clone="SL_EcoRI0062K06"
                        /lab_hosts="E. coli"
                        /clone_lib="Tomato EcoRI BAC Library"
                        /note="Vector: unk; Site_1: EcoRI"

ORIGIN
Query Match          12.6%; Score 37.8; DB 14; Length 950;
Best Local Similarity 54.7%; Pred. No. 46;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      10 AGGCTAAATGCTAGTACGATGAGACTTCGAAATATTTCTTCATTGAAATAAACAG 69
Db      553 ATGGAATAAACTCAATAAAGCGCTCAAACTCAAACTCAATCATATGAGATAAAAG 612
|||||

QY      70 CTAATCTCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTA 129
Db      613 TGACCTTCCAAAGCTTACCACACTAATGAAATTTAATTTGAGCATAGTTACAAAAAATT 672
|||||

QY      130 ATTTTCAAAATAAGAAGT 146
Db      673 ATTACCAAAATCCAAAT 689
|||||

RESULT 41
CZ430880      684 bp mRNA linear EST 25-MAR-2003
LOCUS        606805 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
DEFINITION   CB430880
ACCESSION   CB430880
VERSION     CB430880.1 GI:29207368
KEYWORDS    EST.
SOURCE       Bos taurus (cattle)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 684)
AUTHORS     Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keele,J.W.
TITLE       A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL     Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366

```


Query Match 12.5%; Score 37.6; DB 11; Length 837;
 Best Local Similarity 49.5%; Pred. No. 51;
 Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 43 GAAATATCTTCATTGAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTT 102
 |||||
 Db 202 GAAGCATACATATAGAATAATATAATATCACATTAACTTCTATCATATATTGGTA 143
 |||||
 QY 103 TATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATPAAGAAGTAACACAGAGTGTGCAA 162
 |||||
 Db 142 TAAAGGAAGAAATTAATAATAATAATATACAAATATAAGAAAAATAATAATAGGTA 83
 |||||
 QY 163 CATGCTGTTAAATAACCCGACAACCTCAATCATATAGCTGTAGTAGAGTGCATTCTGC 222
 |||||
 Db 82 AAATACAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 23
 |||||
 QY 223 AAGATCCAGAGTAA 238
 |||||
 Db 22 ATGAATAATAAGTAA 7
 |||||

RESULT 44
 BU355289
 LOCUS 603474334F1 CSEQHN70 Gallus gallus cDNA clone ChEST355d10 5', mRNA
 DEFINITION 991 bp mRNA linear EST 28-NOV-2002
 sequence.
 ACCESSION BU355289
 VERSION BU355289.1 GI:25863290
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 991)
 AUTHORS Boardman, P.E., Sanz-Eguerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.I., Fickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A Comprehensive Collection of Chicken CDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1..991
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST355d10"
 /dev_stage="36"
 /lab_host="DH10B"
 /clone_lib="CSEQHN70"
 /note="Organ: hearts; Vector: pBluescript II KS(+);
 Site 1: EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Query Match 12.5%; Score 37.6; DB 3; Length 991;
 Best Local Similarity 51.8%; Pred. No. 52;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 64 AAACACTAACTCCCAAAATTTGACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAA 123
 |||||
 Db 76 AAACATCGAATCTACTACTATCTTAAAGAGAAAAAGAAAAAGTTTGAAA 135
 |||||
 QY 124 TGCTTAATTTTCAAATPAAGAAGTAACACAGAGTGTTCACACATCTGTATAAATACCCGAC 183
 |||||
 Db 136 TGCTGCATTTACATTCGAAAAAAAACAACTTTTTCACAAATTTTCAACAATTAACAAA 195
 |||||
 QY 184 AAACCTCAATCACTATAGCTGTAGTAGAGTGCATTTCTGCAAGGA 227
 |||||
 Db 196 AATAATTTCAACGCAAAATGGGAGTGTGGGTCTGTTTGGACAGA 239
 |||||

RESULT 45
 R22138/c
 LOCUS yh25D10.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
 DEFINITION IMAGE:130747 3', mRNA sequence.

ACCESSION R22138
 VERSION R22138.1 GI:776919
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 373)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kuchaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevisan, E., Waterston, R., Williamson, A., Wohldmann, P. and
 Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 2451
 High quality sequence stops: 271
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 2451 Std Error: 0.00
 Seq primer: -21m13
 High quality sequence stop: 271.

FEATURES
 source
 1..373
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:543210"
 /db_xref="taxon:9606"
 /clone="IMAGE:130747"
 /sex="Female"
 /dev_stage="placenta obtained at birth (full term)"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares placenta Nb2HP"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 AATCGAAGAAATTCGGCGGCGAGGATTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."


```

ACCESSION CW685077
VERSION CW685077.1 GI:55153667
SOURCE GSS.
ORGANISM Oryza glaberrima (African rice)
Oryza glaberrima
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1. (bases 1 to 704)
AUTHORS Kim,H., Yu,Y., Wisotski,M., Byrne,M., Stum,D., Smart,D., Rao,K.,
Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C.
and Wing,R.
TITLE OMAP
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9395
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0043 row: A column: 15
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
FEATURES
Location/Qualifiers
source 1..704
/organism="Oryza glaberrima"
/mol_type="genomic DNA"
/db_xref="taxon:4538"
/clone="OG BBa0043A15"
/tissue_type="young leaves"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OG BBa"
/note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match 12.5%; Score 37.4; DB 13; Length 704;
Best Local Similarity 53.8%; Pred. No. 56;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 63 TAAACAGCTAACTCCCAAAATTGACCTAACCTAGGGAGTTTATCATGGAAGAAATGTTTAA 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 TATAGAAAAAAATACTAATAATTTTCGAATCAAGATAAATTTATATAGAATATATTTAA 410
QY 123 ATGCTTAATTTTCAAATAAGAGTAACACGAGTGTTCACATGCTGTAAATAACCGA 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
411 CTATTAATTTAAAAAAAATTTGATAATTAATTAATTAATTAATTAATTAATTAATTA 470
QY 183 CAAACTTCAATCACTATAGCTGT 205
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 CAACTAGACGAGTTTAACTTT 493
RESULT 49
BZ047105/c
LOCUS BZ047105 723 bp DNA linear GSS 09-OCT-2002
DEFINITION lxf88e06.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BZ047105
VERSION BZ047105.1 GI:23643266
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1. (bases 1 to 723)
AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Plate: lxf88 row: e column: 06
Seq primer: -21UPpOT forward
Class: shotgun
High quality sequence start: 22
High quality sequence stop: 551.
Location/Qualifiers
source 1..723
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."
ORIGIN
Query Match 12.5%; Score 37.4; DB 11; Length 723;
Best Local Similarity 49.5%; Pred. No. 56;
Matches 95; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 20 CTAAGTGACTACGATCAAGACTTTGAATATCTTCATTTGAAATAACAGCTAACTCCCA 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CAAAGTAATAAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 241
QY 80 AATGTACCTAAGTAGGGAGTTTATCATGGAAGAAATGTTTAAATGCTTAATTTCAAT 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 GAATATACACTCCCATGCTATTTTCTCTAAATATAGATTAATTAATTAATTAATTA 181
QY 140 AGAAGTACCAAGAGTGTTCACATGCTGTAAATAACCCGACAACTTCATCACTAT 199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 AAAAGCAAAAAAATAATTTTCAAANTGACCTTTTGACTTTCAAACCTTAATAATTAT 121
QY 200 AGCTGTAGTAGA 211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 AACTAAATAAA 109
RESULT 50
BZ479314/c
LOCUS BZ479314 797 bp DNA linear GSS 13-DEC-2002
DEFINITION BONHM39TF BO_1.6_2_KB_tot Brassica oleracea genomic clone BONHM39,
genomic survey sequence.
ACCESSION BZ479314
VERSION BZ479314.1 GI:26781712
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1. (bases 1 to 797)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BONHM39TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

```

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TTT

Class: sheared ends.

FEATURES

Source

Class. Shetland Coll.

1. .797

/organism="Brassica oleracea"

```
/mol_type="genomic DNA"
```

```

/strain="TO1000DH3"

```

```
/db xref="taxon:3712"
```

```
/clone="BONHM39"
```

```
/clone lib="BO 1.6 2 KB tot"
```

/note="Vector: pHOS1; Site 1: BstXI; 1.6-2 kb sheared

total DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match	12.5%	Score	37.4;	DB	11;	Length	797;
Best Local Similarity	55.9%	Pred. No.	57;				
Matches	71;	Conservative	0;	Mismatches	56;	Indels	0;
Gaps	0;						
38	GACTTCGAATATATTCTTCATTTTGAATAATAACAGCTAACTCCCAAAATTGTACTCAACTAGGG	97					
746	GACTTCGAAGAGATCATCTATATATAATAGACATCTGATGCCCTTAATTAGGGCTAAGCACTG	687					
98	GAGTTTTATCATGAAGAAATGTTTTAAATGCTTTAATTTTCAAATAAGAAGTAACACAGAGTGT	157					
686	AAATTTGTCTCTAAACCAATTCAGAAAACATTTTGTATGTTGGTGGGAGTGAACAACACCAA	627					
158	TGCAACA	164					
626	TGCAGCA	620					

Search completed: July 17, 2006, 23:22:58
Job time : 2907 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 22:15:53 ; Search time 133 Seconds
(without alignments)
4220.546 Million cell updates/sec

Title: SEQ1-96535C
Perfect score: 300
Sequence: 1 ggcattacaggctaaatgc.....atctgctgattagaagtat 300

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents NA.*

- 1: /EMC_Celerra_SID83/ptodata/2/ina/1 COMB.seq.*
- 2: /EMC_Celerra_SID83/ptodata/2/ina/5 COMB.seq.*
- 3: /EMC_Celerra_SID83/ptodata/2/ina/6A COMB.seq.*
- 4: /EMC_Celerra_SID83/ptodata/2/ina/6B COMB.seq.*
- 5: /EMC_Celerra_SID83/ptodata/2/ina/7 COMB.seq.*
- 6: /EMC_Celerra_SID83/ptodata/2/ina/H COMB.seq.*
- 7: /EMC_Celerra_SID83/ptodata/2/ina/PCUS COMB.seq.*
- 8: /EMC_Celerra_SID83/ptodata/2/ina/PP COMB.seq.*
- 9: /EMC_Celerra_SID83/ptodata/2/ina/RE COMB.seq.*
- 10: /EMC_Celerra_SID83/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
C	1	39.4	13.1	49164	3	US-09-949-016-12985 Sequence 12985, A
C	2	39.4	13.1	49164	3	US-09-949-016-12986 Sequence 12986, A
C	3	39.4	13.1	49164	3	US-09-949-016-13262 Sequence 13262, A
C	4	39.4	13.1	49164	3	US-09-949-016-13263 Sequence 13263, A
C	5	38.4	12.8	192302	3	US-09-949-016-15270 Sequence 15270, A
C	6	36.4	12.1	832	3	US-09-621-976-2813 Sequence 2813, Ap
C	7	35.6	11.9	260247	3	US-09-949-016-13358 Sequence 13358, A
C	8	35.4	11.8	48682	3	US-09-949-016-17236 Sequence 17236, A
C	9	35.4	11.8	48682	3	US-09-949-016-17237 Sequence 17237, A
C	10	35.4	11.8	48682	3	US-09-949-016-17238 Sequence 17238, A
C	11	35.4	11.8	48682	3	US-09-949-016-17239 Sequence 17239, A
C	12	35.4	11.8	51022	3	US-09-949-016-17135 Sequence 17135, A
C	13	35.4	11.8	51022	3	US-09-949-016-17136 Sequence 17136, A
C	14	35.4	11.8	51022	3	US-09-949-016-17137 Sequence 17137, A
C	15	35.4	11.8	51022	3	US-09-949-016-17138 Sequence 17138, A
C	16	35.4	11.8	65415	3	US-09-949-016-16699 Sequence 16699, A
C	17	35.4	11.8	86213	3	US-09-949-016-17240 Sequence 17240, A
C	18	35.4	11.8	86213	3	US-09-949-016-17241 Sequence 17241, A
C	19	35.4	11.8	86213	3	US-09-949-016-17242 Sequence 17242, A
C	20	35.4	11.8	86213	3	US-09-949-016-17243 Sequence 17243, A
C	21	34.6	11.5	821	3	US-08-956-171E-156 Sequence 156, App
C	22	34.6	11.5	821	3	US-08-781-986A-156 Sequence 156, App
C	23	34.2	11.4	38702	3	US-09-949-016-13788 Sequence 13788, A

US-09-949-016-16406
Sequence 16406, A

US-09-949-016-15116
Sequence 15116, A

US-09-949-016-15638
Sequence 15638, A

US-09-543-681A-714
Sequence 714, App

US-09-949-016-15744
Sequence 15744, A

US-09-949-016-16600
Sequence 16600, A

US-09-949-016-53968
Sequence 53968, A

US-09-949-016-12025
Sequence 12025, A

US-09-949-016-13302
Sequence 13302, A

US-09-621-976-2813
Sequence 2813, Ap

US-08-929-329-3
Sequence 3, Appli

US-09-662-254B-24
Sequence 24, Appli

US-09-949-016-13185
Sequence 13185, A

US-09-949-016-13186
Sequence 13186, A

US-09-949-016-13187
Sequence 13187, A

US-09-949-016-13188
Sequence 13188, A

US-09-949-016-92131
Sequence 92131, A

US-09-949-016-16290
Sequence 16290, A

US-09-949-016-13659
Sequence 13659, A

US-09-949-016-12582
Sequence 12582, A

US-09-596-002-40
Sequence 13176, A

US-09-949-016-14382
Sequence 40, Appli

US-09-949-016-17030
Sequence 17030, A

US-09-248-796A-687
Sequence 687, App

US-09-949-016-12723
Sequence 12723, A

US-09-949-016-16252
Sequence 16252, A

US-09-617-594A-3
Sequence 3, Appli

US-10-209-507-3
Sequence 3, Appli

US-09-949-016-179050
Sequence 179050, A

US-09-949-016-179051
Sequence 179051, A

US-09-949-016-196473
Sequence 196473, A

US-09-487-558B-305
Sequence 305, App

US-09-949-016-16908
Sequence 16908, A

US-09-762-311-2
Sequence 2, Appli

US-09-949-016-16151
Sequence 16151, A

US-09-949-016-12896
Sequence 12896, A

US-09-949-016-17305
Sequence 17305, A

US-08-916-421B-1
Sequence 1, Appli

US-09-513-999C-23091
Sequence 23091, A

US-09-710-279-1651
Sequence 1651, Ap

US-09-134-001C-1237
Sequence 1237, Ap

US-09-710-279-4273
Sequence 4273, Ap

US-09-949-016-13703
Sequence 13703, A

US-09-949-016-16179
Sequence 16179, A

US-09-949-016-16180
Sequence 16180, A

US-09-949-016-16297
Sequence 16297, A

US-09-949-016-35532
Sequence 35532, A

US-09-949-016-62973
Sequence 62973, A

US-09-543-681A-903
Sequence 903, App

US-09-949-016-910
Sequence 910, App

US-09-949-016-1818
Sequence 1818, Ap

US-08-213-419B-3
Sequence 3, Appli

US-09-949-016-12652
Sequence 12652, A

US-09-949-016-13560
Sequence 13560, A

US-09-949-016-14320
Sequence 14320, A

US-09-949-016-16445
Sequence 16445, A

US-09-531-120-202
Sequence 202, App

US-09-949-016-14720
Sequence 14720, A

US-09-949-016-196410
Sequence 196410, A

US-09-949-016-196412
Sequence 196412, A

US-09-248-796A-5116
Sequence 5116, Ap

US-09-949-016-4723
Sequence 4723, Ap

US-09-936-271C-21
Sequence 21, Appli

US-09-949-016-16412
Sequence 16412, A

US-09-949-016-13451
Sequence 13451, A

US-09-276-625-12
Sequence 12, Appli

US-10-054-665A-12
Sequence 12, Appli


```
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13262
LENGTH: 49164
TYPE: DNA
ORGANISM: Human
US-09-949-016-13262

Query Match      13.1%; Score 39.4; DB 3; Length 49164;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAATAAACAGCTAACTCC 77
DB 24145 TACATAGTGAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAATTAA 24086
QY 78 CAAATTGACTTAAGTACAGGAGTTTATCATGAGAGAAATGTTAAATGCTTAATTTCAA 137
DB 24085 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGATAAAACAAATAATGTAA 24026
QY 138 ATAAGAGTAAACACAGAG 154
DB 24025 CTAATAAATAAGAGAG 24009

US-09-949-016-13263/c
Sequence 13263, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13263
LENGTH: 49164
TYPE: DNA
ORGANISM: Human
US-09-949-016-13263

Query Match      13.1%; Score 39.4; DB 3; Length 49164;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAATAAACAGCTAACTCC 77
DB 24145 TACATAGTGAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAATTAA 24086
QY 78 CAAATTGACTTAAGTACAGGAGTTTATCATGAGAGAAATGTTAAATGCTTAATTTCAA 137
DB 24085 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGATAAAACAAATAATGTAA 24026
QY 138 ATAAGAGTAAACACAGAG 154
DB 24025 CTAATAAATAAGAGAG 24009

RESULT 4
US-09-949-016-13263/c
Sequence 13263, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13263
LENGTH: 49164
TYPE: DNA
ORGANISM: Human
US-09-949-016-13263

Query Match      13.1%; Score 39.4; DB 3; Length 49164;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAATAAACAGCTAACTCC 77
DB 24145 TACATAGTGAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAATTAA 24086
QY 78 CAAATTGACTTAAGTACAGGAGTTTATCATGAGAGAAATGTTAAATGCTTAATTTCAA 137
DB 24085 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGATAAAACAAATAATGTAA 24026
QY 138 ATAAGAGTAAACACAGAG 154
DB 24025 CTAATAAATAAGAGAG 24009

RESULT 5
US-09-949-016-15270/c
Sequence 15270, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15270
LENGTH: 192302
TYPE: DNA
ORGANISM: Human
NAME/KEY: misc feature
FEATURE:
LOCATION: (1)-(192302)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15270

Query Match      12.8%; Score 38.4; DB 3; Length 192302;
Best Local Similarity 51.8%; Pred. No. 0.47;
Matches 87; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 74 CTCCTCAATTTGCTACCTAAGTGGGAGTTTATCATGAGAGAAATGTTTAAATGCTTAATTT 133
DB 66927 CTACTCCATTTGGATCTAATCTTTGAGATACCAAGAAATATATTTGAGCATATCTCTATAGG 66868
QY 134 TCAATAGAGAGTAACAGAGTGTTCACACATGCTGTAAATACCCGACAACTTCAAT 193
DB 66867 TTAGTACGGAGTTACATCTGACATATTAATTTGAGTGTACATAAGAAAGTAGACTGAAAT 66808
QY 194 CACTATAGCTGTAGTAGAGTGCATTCTGCAAGGATCCAGAGTAACCA 241
DB 66807 ACATACAGCATTTGAGTATAGAGTTTCAAGGGTTCAAGGGTAGCAA 66760

RESULT 6
US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PE2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
```

[illegible]

```

RESULT 10
US-09-949-016-17238/c
; Sequence 17238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17238
; LENGTH: 48682
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17238

Query Match      11.8%; Score 35.4; DB 3; Length 48682;
Best Local Similarity 56.4%; Pred. No. 2.4;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAAACATCTTATATAAACAGCTAAAGTATTAACTT 14446
QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGAATTACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 11
US-09-949-016-17239/c
; Sequence 17239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17239
; LENGTH: 48682
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17239

Query Match      11.8%; Score 35.4; DB 3; Length 48682;
Best Local Similarity 56.4%; Pred. No. 2.4;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAAACATCTTATATAAACAGCTAAAGTATTAACTT 14446
QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGAATTACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 12
US-09-949-016-17135/c
; Sequence 17135, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17135
; LENGTH: 51022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17135

Query Match      11.8%; Score 35.4; DB 3; Length 51022;
Best Local Similarity 56.4%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAAACATCTTATATAAACAGCTAAAGTATTAACTT 14446
QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGAATTACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 13
US-09-949-016-17136/c
; Sequence 17136, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17136
; LENGTH: 51022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17136

Query Match      11.8%; Score 35.4; DB 3; Length 51022;
Best Local Similarity 56.4%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAAACATCTTATATAAACAGCTAAAGTATTAACTT 14446
QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGAATTACACTTGAAATGTATAATGATATCTGATCAATTA 14389
```

Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACATCTTATATAAAGCAGCTTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGATTACACTTTGAAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 14
 US-09-949-016-17137/c
 ; Sequence 17137, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17137
 ; LENGTH: 51022
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-17137

Query Match 11.8%; Score 35.4; DB 3; Length 51022;
 Best Local Similarity 56.4%; Pred. No. 2.5;
 Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 24 GTGACTACGAGTGAAGCTTGAATATTTCTTCATTTGAATAAAGCAGCTTAACCTCCCAATT 83
 Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACATCTTATATAAAGCAGCTTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGATTACACTTTGAAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 15
 US-09-949-016-17138/c
 ; Sequence 17138, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17138
 ; LENGTH: 51022
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-17138
 Query Match 11.8%; Score 35.4; DB 3; Length 51022;
 Best Local Similarity 56.4%; Pred. No. 2.5;
 Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGAGTGAAGCTTGAATATTTCTTCATTTGAATAAAGCAGCTTAACCTCCCAATT 83
 Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACATCTTATATAAAGCAGCTTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGATTACACTTTGAAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 16
 US-09-949-016-16699
 ; Sequence 16699, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16699
 ; LENGTH: 65415
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-16699

Query Match 11.8%; Score 35.4; DB 3; Length 65415;
 Best Local Similarity 51.6%; Pred. No. 2.7;
 Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 5 TTGACAGGCTAAATGCTAGTACTAGATGAAGACTTGAAATATTTCTTCATTTGAAATA 64
 Db 739 TTCAACAATGATTCAAAAGTAAAGAACTAGAGAAAAAGATAAAATTTTATATAAGAACTA 798
 QY 65 ACAGCTAACTCCCAATTTGTTACTTAAGTGGGAGTTTATCATGAAGAAATGTTTAAAT 124
 Db 799 AATCACCATTTCACAGAGGCCACGACTGTGGAACTGCAGAGACGACACTGGGAAT 858
 QY 125 GCTTAATTTTCAATAAAGAAAGTAAACAGAGCTGTTGCA 161
 Db 859 GCATTATTACAAATGAATAATTTTAAACAAATCTTTCA 895

RESULT 17
 US-09-949-016-17240/c
 ; Sequence 17240, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17240
 ; LENGTH: 86213

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17240

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAGTATTAACTT 14446

QY 84 GTACTTAAGTGGGAGTTTATCATGAAGAATGTTTAAATGCTTAAATTTCAATA 140
DB 14445 TATCTCTTTAAGCTAGAAATTACACTTGAAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 18
US-09-949-016-17241/c
; Sequence 17241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17241
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17241

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAGTATTAACTT 14446

QY 84 GTACTTAAGTGGGAGTTTATCATGAAGAATGTTTAAATGCTTAAATTTCAATA 140
DB 14445 TATCTCTTTAAGCTAGAAATTACACTTGAAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 19
US-09-949-016-17242/c
; Sequence 17242, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17242
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17242

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAGTATTAACTT 14446

QY 84 GTACTTAAGTGGGAGTTTATCATGAAGAATGTTTAAATGCTTAAATTTCAATA 140
DB 14445 TATCTCTTTAAGCTAGAAATTACACTTGAAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 20
US-09-949-016-17243/c
; Sequence 17243, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17243
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17243

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAGTATTAACTT 14446

QY 84 GTACTTAAGTGGGAGTTTATCATGAAGAATGTTTAAATGCTTAAATTTCAATA 140
DB 14445 TATCTCTTTAAGCTAGAAATTACACTTGAAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 21
US-08-956-171E-156/c
; Sequence 156, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
```

```
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-956-171E-156
Query Match 11.5%; Score 34.6; DB 3; Length 821;
Best Local Similarity 56.6%; Pred. No. 1.2;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 105 TCATGAAGAAATGTTAAATGCTTAAATTTTCAATAAGAAGTAAACCAAGAGTGTTCGAACA 164
DB 135 TTATTATTATAAATTTCCAACTTAGCTTAGATTGAGAAATATAAAAGTATTAAACA 76
QY 165 TGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCAT 217
DB 75 CCCAGAAATATTGCAACACATATTACAATCACCATGTTTTTCTCTAATTGAAT 23
RESULT 22
US-08-781-986A-156/c
; Sequence 156, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Scaphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781.986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-956-171E-156
Query Match 11.5%; Score 34.6; DB 3; Length 821;
Best Local Similarity 56.6%; Pred. No. 1.2;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 105 TCATGAAGAAATGTTAAATGCTTAAATTTTCAATAAGAAGTAAACCAAGAGTGTTCGAACA 164
DB 135 TTATTATTATAAATTTCCAACTTAGCTTAGATTGAGAAATATAAAAGTATTAAACA 76
QY 165 TGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCAT 217
DB 75 CCCAGAAATATTGCAACACATATTACAATCACCATGTTTTTCTCTAATTGAAT 23
RESULT 23
US-09-949-016-13788/c
; Sequence 13788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13788
; LENGTH: 38702
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38702)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13788
Query Match 11.4%; Score 34.2; DB 3; Length 38702;
Best Local Similarity 50.9%; Pred. No. 5.2;
Matches 81; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 108 TGAAGAAATGTTAAATGCTTAAATTTTCAATAAGAAGTAAACCAAGAGTGTTCGAACATGC 167
DB 35855 TGAAGTGTAGCGAAGCTGCTGGAAGTTTGAATGAGATCTCTTAACAGCTCTGTATACCT 35796
QY 168 TGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTCTGCAAGGA 227
DB 35795 AACTAAAAAACCTGCCAAACCATATTAACTATTCTGGAGATTCTTCAATCTTGCACCTT 35736
QY 228 TCCAGAGTAAACCAAGTATTGGAATGCAATGTTGAAC 266
```

```
DB 35735 GCATTAGAAAGTTCTGAATGGCTTAAGAAAGGATGTAC 35697

RESULT 24
US-09-949-016-16406
; Sequence 16406, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16406
; LENGTH: 68778
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(68778)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16406

Query Match 11.3%; Score 34; DB 3; Length 68778;
Best Local Similarity 52.9%; Pred. No. 7.1;
Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 39 ACTTGAATATCTTCATTTGAAATAACACCTAACTCCCAATGTACCTAGGAGTGTATCATGAAGAAA 115
DB 38112 ACTTGAATATCTTCGATTTTAAATAACAGATATATACACAAATGTATGTGTGTA 38171
QY 99 AGTTTATCATCAAGAAATGTTTAAATGCTTAAATTTTCAAAATAAGAGTAACACGAGTGT 158
DB 38172 TGATCAATAAAGAACAGAGTAGTTTATAGATATCTGTAATACATCTGTTGTTGT 38231
QY 159 GCAACATCTGTTAAATA 176
DB 38232 GGATCAATTAATAAATA 38249

RESULT 25
US-09-949-016-15116
; Sequence 15116, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15116
; LENGTH: 83462
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(83462)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15116

Query Match 11.3%; Score 33.8; DB 3; Length 69764;
Best Local Similarity 55.6%; Pred. No. 8.2;
Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 33 ATGAAGACTTGAATAATTTTCATTTGAAATAAAGAGTAACCTCCCAATTTGTACCTAAC 92
DB 29557 ATGAAGCTAGAACATTTCTAAATGTCCACAAAATGCTTTTATTAATTCGGTATAC 29616
QY 93 TAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTTAAATGCTTTTCAAAATAAGAGTAAC 149
DB 29617 CCCTGAAGTGAATCATTAGGTAACTTTAGAGTTTAAATTTATTAAGAGTCTTTTAAATAC 29673

RESULT 27
US-09-543-681A-714/c
; Sequence 714, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
```



```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13186
; LENGTH: 251769
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13186

Query Match      11.0%; Score 33; DB 3; Length 251769;
Best Local Similarity 57.1%; Pred. No. 21;
Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 161 AACATGCTGTTAAATAACCCGACAACTTCAATCACTAGCTAGTAGAGTGCAATCT 220
DB 171714 AACATCCTTTTCATATATTTCTCGGCACCTTCTATCAGTATATATGCGAGTAGAGTACTAG 171655

QY 221 GCAAGGATCCAGAGTAACCAAGTATTTTGGAAATGCAATGTTGAA 265
DB 171654 ACATGGACCTGGAGGTCAAGAATAGGTGATTTTAAATTTTAA 171610

RESULT 38
US-09-949-016-13187/c
; Sequence 13187..Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13187
; LENGTH: 266748
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13187

Query Match      11.0%; Score 33; DB 3; Length 266748;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 161 AACATGCTGTTAAATAACCCGACAACTTCAATCACTAGCTAGTAGAGTGCAATCT 220
DB 171679 AACATCCTTTTCATATATTTCTCGGCACCTTCTATCAGTATATATGCGAGTAGAGTACTAG 171620

QY 221 GCAAGGATCCAGAGTAACCAAGTATTTTGGAAATGCAATGTTGAA 265
DB 171619 ACATGGACCTGGAGGTCAAGAATAGGTGATTTTAAATTTTAA 171575

RESULT 39
US-09-949-016-13188/c
; Sequence 13188..Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13188
; LENGTH: 266748
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13188

Query Match      11.0%; Score 33; DB 3; Length 266748;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 161 AACATGCTGTTAAATAACCCGACAACTTCAATCACTAGCTAGTAGAGTGCAATCT 220
DB 171679 AACATCCTTTTCATATATTTCTCGGCACCTTCTATCAGTATATATGCGAGTAGAGTACTAG 171620

QY 221 GCAAGGATCCAGAGTAACCAAGTATTTTGGAAATGCAATGTTGAA 265
DB 171619 ACATGGACCTGGAGGTCAAGAATAGGTGATTTTAAATTTTAA 171575

RESULT 40
US-09-949-016-92131/c
; Sequence 92131..Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92131
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-92131

Query Match      10.9%; Score 32.8; DB 3; Length 601;
Best Local Similarity 49.1%; Pred. No. 3.7;
Matches 114; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 30 ACGATGAAGACTTGAATATTTCTTCAATTTGAAATAAACAGCTAACTCCCAATTTGACCT 89
DB 308 ACAATCASCACCTTGACTTTTACTTCA-GGAAATAAAGTTTTTACTTGGGAACAGATATTT 250

QY 90 AACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAGTAAAC 149
DB 249 TAGCATAAGAACTTAGCATTAAGAACTAAATGATTAGTATTTCCAAAGGTACAAAAT 190

QY 150 CAGAGTGTTCACATGCTGTTTAAATAACCCGACAACTTCAATCACTATATAGCTGTAGTA 209
DB 189 CATATTTGGCCAAATTCCTTCAAAAGCCAGCTACCACCTACTACAGACAGATCTGGGTGA 130

QY 210 GAGTGCATTTCTGCAAGGATCCAGAGTAACCAAGTATTTTGGAAATGCAATGCT 261
DB 129 CGATGCGTGCTTATTTGCAACAGTCAACTTCTGGACCTTAGAATGTGCCCTCT 78
```

```

RESULT 41
US-09-949-016-16290
; Sequence 16290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16290
; LENGTH: 10783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16290

Query Match 10.9%; Score 32.8; DB 3; Length 10783;
Best Local Similarity 56.5%; Pred. No. 9.1;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 27 ACTACAGTGAAGACTGAAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTA 86
DB 7795 ACAACAGTCAATAAGAGAAAGTACAAAATTCAAAGTAAGAAATTAACCTTGGATTTGAA 7854

QY 87 CCTAAGTGGGAGTTTATCATGAGAAATCTTTAAATGCTTAATTTT 134
DB 7855 GCTTACTTTTCTAGATTATGTTTCAGAGATCTTTGACATTTTAACTTT 7902

RESULT 42
US-09-949-016-13659
; Sequence 13659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13659
; LENGTH: 23105
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13659

Query Match 10.9%; Score 32.8; DB 3; Length 23105;
Best Local Similarity 49.4%; Pred. No. 12;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 129 AATTTTCAATAAGAGTAAACAGAGTGTTCGAACATCTGTTAAATAACCCGACAACT 188
DB 7977 AATTTTCAATGATGATTTACCAATAGTTTCTAAATACATTTTAATAAGTCCCTAGACT 8036

RESULT 43
US-09-949-016-12582
; Sequence 12582, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12582
; LENGTH: 23108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12582

Query Match 10.9%; Score 32.8; DB 3; Length 23108;
Best Local Similarity 49.4%; Pred. No. 12;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 129 AATTTTCAATAAGAGTAAACAGAGTGTTCGAACATCTGTTAAATAACCCGACAACT 188
DB 7980 AATTTTCAATGATGATTTTACCATAGTTTCTAAATACATTTTAATAAGTCCCTAGACT 8039

QY 189 TCAATCACTATAGCTGTAGTAGAGTGCATCTCTCAAGGATCCAGAGTAACCACTATTTT 248
DB 8040 TTGAACCTCTTTCCCTCTGGAGCAGACCTTGGGCGAAGGATGTGGAGCATGTAGTTTATT 8099

QY 249 GGAATGCAATGTTGAACCGACCATACTAATTAATCTGCTGATTAGGAAGTAT 300
DB 8100 TGGAGTGTATCTTAGGAAGAGGGATTAAGGAGGAACCTTAGGAGGACTAT 8151

RESULT 44
US-09-949-016-13176
; Sequence 13176, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13176
; LENGTH: 36577
; TYPE: DNA

```

```

QY 189 TCAATCACTATAGCTGTAGTAGAGTGCATCTCTCAAGGATCCAGAGTAACCACTATTTT 248
DB 8037 TTGAACCTCTTTCCCTCTGGAGCAGACCTTGGGCGAAGGATGTGGAGCATGTAGTTTATT 8096

QY 249 GGAATGCAATGTTGAACCGACCATACTAATTAATCTGCTGATTAGGAAGTAT 300
DB 8097 TGGAGTGTATCTTAGGAAGAGGGATTAAGGAGGAACCTTAGGAGGACTAT 8148

RESULT 43
US-09-949-016-12582
; Sequence 12582, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12582
; LENGTH: 23108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12582

Query Match 10.9%; Score 32.8; DB 3; Length 23108;
Best Local Similarity 49.4%; Pred. No. 12;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 129 AATTTTCAATAAGAGTAAACAGAGTGTTCGAACATCTGTTAAATAACCCGACAACT 188
DB 7980 AATTTTCAATGATGATTTTACCATAGTTTCTAAATACATTTTAATAAGTCCCTAGACT 8039

QY 189 TCAATCACTATAGCTGTAGTAGAGTGCATCTCTCAAGGATCCAGAGTAACCACTATTTT 248
DB 8040 TTGAACCTCTTTCCCTCTGGAGCAGACCTTGGGCGAAGGATGTGGAGCATGTAGTTTATT 8099

QY 249 GGAATGCAATGTTGAACCGACCATACTAATTAATCTGCTGATTAGGAAGTAT 300
DB 8100 TGGAGTGTATCTTAGGAAGAGGGATTAAGGAGGAACCTTAGGAGGACTAT 8151

RESULT 44
US-09-949-016-13176
; Sequence 13176, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13176
; LENGTH: 36577
; TYPE: DNA

```

```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36577)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13176

Query Match          10.9%; Score 32.8; DB 3; Length 36577;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 33 ATGAAGACTTGAAATATTTCTTCATTTGAAATAAAGACGCTAACTCCCAAAATTGTACCTAAC 92
DB 21634 AGGAAGATGCATATCTCTGCTTTGGGAATAAGAACCAAACTGGAATCCAGGTGGC 21693

QY 93 TAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACCAAG 152
DB 21694 TAGTGGTGTAGACTGTTAAGTAATGTTTAAAGAGTGTGAGTTTACGATAATAGATTGAAT 21753

QY 153 AGTGTGCAACATGCTGTTTAATAATACCCGACAAACT 188
DB 21754 TGATTAGATACAAATGGTTTAATAATTTTAAAAAAT 21789

RESULT 45
US-09-596-002-40 :
; Sequence 40, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 119211
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 40
; PUBLICATION INFORMATION:
US-09-596-002-40

Query Match          10.9%; Score 32.8; DB 3; Length 119211;
Best Local Similarity 52.1%; Pred. No. 19;
Matches 73; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1 GGCATTCACAGCGCTAAATGCTTAAGTGACTAGATGACATGATGAGTGCAGCAATCAATAATCACTTGATCTTGCCATGTTTAA 108095
DB 108036 GGGCTGGATAAAGTGAAGTGGTGCAGCAATCAATAATCACTTGATCTTGCCATGTTTAA 108095

QY 61 AATAAAGCTTAACCTCCCAATTTGTACCTACTAGGGGAGTTTATCATGACAGAAATGTTT 120
DB 108096 AATCAAAATTTTATTCCTCAATATGCGATATAATTTATAAATTTATGTCATAATACAAAGTTT 108155

QY 121 AAATGCTTTAATTTTCAATA 140
DB 108156 TAAAGTGATATTTCAATA 108175

RESULT 46
US-09-949-016-14382/c
; Sequence 14382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14382
; LENGTH: 192956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(192956)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14382

Query Match          10.9%; Score 32.8; DB 3; Length 192956;
Best Local Similarity 49.1%; Pred. No. 22;
Matches 114; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 30 ACGATGAAGACTTGAAATATTTCTTCATTTGAAATAAAGACGCTAACTCCCAAAATTGTACCT 89
DB 4134 ACAATCAGCACTTGACTTTTACTTCA-GGAAATAAAGTTTACTTTGGGAACAGATATTT 4076

QY 90 AACTAGGGAGCTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATGAAGTAAC 149
DB 4075 TAGCATGAAGACTTAGCATAGAACAACAACTAAATGATTAGTATTATCCAAAGGTACAAAAT 4016

QY 150 CAGAGTGTGCAACATGCTGTTAAATAACCGACAACTTCACTCATCTATAGCTGTAGTA 209
DB 4015 CATATTTGGCCAAAATCCTTCAAAAGCCAAAGTACCCTACTACTACAGACAGATCTGGTGA 3956

QY 210 GAGTGCATTTCTGCAAGGATCCAGAGTAACAGTATTTTGGAAATGCAATGT 261
DB 3955 CGATGCGTCTTATTGCAACAGTGAACCTTGGACCTTAGAATGTGCCCTGT 3904

RESULT 47
US-10-002-623-923/c
; Sequence 923, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OEFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 923
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-923

Query Match          10.9%; Score 32.6; DB 3; Length 456;
Best Local Similarity 60.9%; Pred. No. 3.9;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 55 ATTTGAAATAAACAGCTAACTCCCAAAATTTGACCTAACTAGGGAGTTTATCATCAAGAA 114

```

```

Db 297 ATCCAAAGAAAAAATAAATTTTGTAAATATAGACTACTTTATGTTGTAATA 238
Qy 115 ATGTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

RESULT 48
US-10-002-623-926/c
; Sequence 926, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OEFNER, PETER J.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 926
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-926

Query Match 10.9%; Score 32.6; DB 3; Length 456;
Best Local Similarity 60.9%; Pred. No. 3.9;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 55 ATTTGAATAACAGCTAACTCCCAATTTGTACCTACTAGGGGAGTTTATCATGAAGAA 114
Db 297 ATCCAAAGAAAAAATAAATTTTGTAAATATAGACTACTTTATGTTGTAATA 238

Qy 115 ATGTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

RESULT 49
US-10-002-623-929/c
; Sequence 929, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OEFNER, PETER J.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 929
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-929

Query Match 10.9%; Score 32.6; DB 3; Length 623;
Best Local Similarity 60.9%; Pred. No. 4.3;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 55 ATTTGAATAACAGCTAACTCCCAATTTGTACCTACTAGGGGAGTTTATCATGAAGAA 114
Db 297 ATCCAAAGAAAAAATAAATTTTGTAAATATAGACTACTTTATGTTGTAATA 238

```

```

Qy 115 ATGTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

```

```

RESULT 50
US-09-809-545A-64/c
; Sequence 64, Application US/09809545A
; Patent No. 6800455
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence W.
; APPLICANT: White, R. Tyler
; TITLE OF INVENTION: SECRETED FACTORS
; FILE REFERENCE: SCIOS.017A
; CURRENT APPLICATION NUMBER: US/09/809,545A
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (0)...(0)
; OTHER INFORMATION: n = A, T, C, or G
US-09-809-545A-64

```

```

Query Match 10.9%; Score 32.6; DB 3; Length 819;
Best Local Similarity 47.0%; Pred. No. 4.7;
Matches 101; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 39 ACTTGAATATCTTCTTCAATTTGAATAAACAAGCTAACTCCCAATTTGTACCTAACTAGGGG 98
Db 371 ACTTGAATATACCACACTTTTAAAAAGGCAACAAGTGACCAGATTTAAATTTAAGATAGT 312

Qy 99 AGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAGTAACCAGAGTGT 158
Db 311 ATATTGCAAAATACATCAGCATATATATGTCGAGCACAAGCAAGTAGTTTCCATTCC 252

Qy 159 GCACATGCTGTTAAATAACCCGACAAACITTCATCACTATAGCTAGTAGTAGTGCAATT 218
Db 251 GAAGGATACCTTTTAACTTCTAATAAGCAACCGTTCCTCAATTTTGTGTTGTTCAAGT 192

Qy 219 CTGCAAGGATCCAGAGTAACCCAGATTTTGGAAA 253
Db 191 CTAATATATACAAAGAGTGATACATTTTACACA 157

```

Search completed: July 18, 2006, 01:05:41
Job time : 140 secs


```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 22:44:13 ; Search time 1462 Seconds
        (without alignments)
        2521.403 Million cell updates/sec

Title:      SEQ1-96535C
Perfect score: 300
Sequence: 1 ggcattgacggctaaatgc.....atctgctgattaggaagtat 300

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 19892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Published Applications NA Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description
-----
1 299.6 100.0 447894 10 US-10-723-670-1 Sequence 1, Appli
2 298 99.3 614 4 US-09-925-065A-428890 Sequence 428890,
3 298 99.3 614 5 US-09-925-065A-428890 Sequence 428890,
4 222 74.0 627 6 US-10-027-632-116164 Sequence 116164,
5 222 74.0 627 6 US-10-027-632-116165 Sequence 116165,
6 222 74.0 627 7 US-10-027-632-116164 Sequence 116164,
7 222 74.0 627 7 US-10-027-632-116165 Sequence 116165,
8 40.6 13.5 201 10 US-10-995-561-132777 Sequence 132777, A
9 40.6 13.5 201 10 US-10-995-561-132777 Sequence 132777, A
10 40.6 13.5 588 4 US-09-925-065A-349719 Sequence 349719,
11 40.6 13.5 588 5 US-09-925-065A-349719 Sequence 349719,
12 40.6 13.5 599 12 US-10-301-480-421781 Sequence 421781,
13 40.6 13.5 599 12 US-10-301-480-421782 Sequence 421782,
14 40.6 13.5 57198 10 US-10-995-561-132663 Sequence 132663, A
15 40.6 13.5 173995 10 US-10-995-561-132777 Sequence 132777, A
16 39.6 13.2 483 12 US-10-301-480-21248 Sequence 21248, A
17 39.6 13.2 483 12 US-10-301-480-634657 Sequence 634657,

Sequence 93552, A
Sequence 305099, A
Sequence 93552, A
Sequence 305099, A
Sequence 349720, A
Sequence 349722, A
Sequence 349722, A
Sequence 421780, A
Sequence 421782, A
Sequence 1035189, A
Sequence 1035191, A
Sequence 349721, A
Sequence 421781, A
Sequence 1035190, A
Sequence 82, Appl
Sequence 331, Appl
Sequence 21832, A
Sequence 22160, A
Sequence 22166, A
Sequence 23173, A
Sequence 24502, A
Sequence 24579, A
Sequence 24590, A
Sequence 25046, A
Sequence 27679, A
Sequence 28021, A
Sequence 28027, A
Sequence 29041, A
Sequence 86622, A
Sequence 86623, A
Sequence 316419, A
Sequence 316420, A
Sequence 316420, A
Sequence 86622, A
Sequence 86623, A
Sequence 316419, A
Sequence 316420, A
Sequence 18060, A
Sequence 631469, A
Sequence 248299, A
Sequence 248300, A
Sequence 861708, A
Sequence 861709, A
Sequence 154831, A
Sequence 154832, A
Sequence 154831, A
Sequence 154832, A
Sequence 154832, A
Sequence 203218, A
Sequence 816627, A
Sequence 102887, A
Sequence 102887, A
Sequence 18061, A
Sequence 631470, A
Sequence 929082, A
Sequence 950049, A
Sequence 950049, A
Sequence 929082, A
Sequence 950049, A
Sequence 16, Appl
Sequence 50, Appl
Sequence 52, Appl
Sequence 54, Appl
Sequence 56, Appl
Sequence 58, Appl
Sequence 60, Appl
Sequence 62, Appl
Sequence 64, Appl
Sequence 66, Appl
Sequence 68, Appl
Sequence 70, Appl
Sequence 72, Appl
Sequence 74, Appl
Sequence 76, Appl

```



```
US-09-925-065A-428890
Query Match          99.3%; Score 298; DB 4; Length 614;
Best Local Similarity 99.3%; Pred. No. 2.2e-70;
Matches 298; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 60
Db 105 GGCATTGACAGGCTAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 164

Qy 61 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 120
Db 165 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 224

Qy 121 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 225 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 284

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTCAAGGATCCAGAGTAACC 240
Db 285 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTCAAGGATCCAGAGTAACC 344

Qy 241 AGTATTTTGGAAATGCAATGTTGAACCGACCATCTAATTAATCTGCTGATTAGGAAGTAT 300
Db 345 AGTATTTTGGAAATGCAATGTTGAACCGACCATCTAATTAATCTGCTGATTAGGAAGTAT 404

RESULT 3
US-09-925-065A-428890
; Sequence 428890; Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428890
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-428890

Query Match          99.3%; Score 298; DB 5; Length 614;
Best Local Similarity 99.3%; Pred. No. 2.2e-70;
Matches 298; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 60
Db 105 GGCATTGACAGGCTAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 164

Qy 61 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 120
Db 165 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 224

Qy 121 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 225 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 284

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTCAAGGATCCAGAGTAACC 240
Db 285 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTCAAGGATCCAGAGTAACC 344

US-09-925-065A-428890
Query Match          74.0%; Score 222; DB 6; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 60
Db 224 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGTGTGAAATATTTCTTCATTGA 165

Qy 61 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 120
Db 164 AATAAAGAGTAACTCCCAATTTGCTAATCTAAGGAGTGTTCATGAAGAAGTATTT 105

Qy 121 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 104 AAATGCTTAATTTTCAATTAAGAAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 45

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTGCAA 224
Db 44 GACAACTTCAATCACTATAGTCTGTAGTACGATGCTTCTGCAA 1

RESULT 5
US-10-027-632-116164/c
; Sequence 116164; Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
```


QY 181 GACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCGCA 224
DB |||||
44 GACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCGCA 1

RESULT 8
US-10-995-561-33431/c
; Sequence 33431, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33431
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33431

Query Match 13.5%; Score 40.6; DB 10; Length 201;
Best Local Similarity 55.5%; Pred. No. 1;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;
QY 18 TGCTAAGTACGATGACGAACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
188 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 129
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
DB |||||
128 CAGAGTTATTAAAGTAAAGAGCNAAYAAAGAGAAATGATGAACAAATATGTAA 69
QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
68 CTAAGAAATAAAGAGAG 52

RESULT 9
US-10-995-561-37932/c
; Sequence 37932, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37932
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-37932

Query Match 13.5%; Score 40.6; DB 10; Length 201;
Best Local Similarity 55.5%; Pred. No. 1;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;
QY 18 TGCTAAGTACGATGACGAACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
188 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 129
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
DB |||||
128 CAGAGTTATTAAAGTAAAGAGCNAAYAAAGAGAAATGATGAACAAATATGTAA 69

QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
68 CTAAGAAATAAAGAGAG 52

RESULT 10
US-09-925-065A-349719
; Sequence 349719, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349719
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349719

Query Match 13.5%; Score 40.6; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 1.6;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;
QY 18 TGCTAAGTACGATGACGAACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 168
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
DB |||||
169 CAGAGTTATTAAAGTAAAGAGCNAAYAAAGAGAAATGATGAACAAATATGTAA 228
QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
229 CTAAGAAATAAAGAGAG 245

RESULT 11
US-09-925-065A-349719
; Sequence 349719, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

```
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349719
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349719

Query Match      13.5%; Score 40.6; DB 5; Length 588;
Best Local Similarity 55.5%; Pred. No. 1.6;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

Qy 18 TGCTAAGTGACTACGATGAAGACCTTGAATATCTTCATTGAAATAAACAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAAAAATTAA 168

Qy 78 CAAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CAGAAGTTATTAAAGTAAAGAGCAAAAYAAAGAGAAGTAAGTAAACAAATATGTAA 228

Qy 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 CTAAAAAATAAGAGAG 245

RESULT 12
US-10-301-480-421779
; Sequence 421779, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 421779
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-421779

Query Match      13.5%; Score 40.6; DB 12; Length 599;
Best Local Similarity 55.5%; Pred. No. 1.7;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

Qy 18 TGCTAAGTGACTACGATGAAGACCTTGAATATCTTCATTGAAATAAACAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAAAAATTAA 168

Qy 78 CAAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CAGAAGTTATTAAAGTAAAGAGCAAAAYAAAGAGAAGTAAGTAAACAAATATGTAA 228

Qy 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 CTAAAAAATAAGAGAG 245

RESULT 13
US-10-301-480-1035188
; Sequence 1035188, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1035188
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1035188

Query Match      13.5%; Score 40.6; DB 10; Length 57198;
Best Local Similarity 55.5%; Pred. No. 11;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

Qy 18 TGCTAAGTGACTACGATGAAGACCTTGAATATCTTCATTGAAATAAACAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28145 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAAAAATTAA 28086

Qy 78 CAAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28085 CAGAAGTTATTAAAGTAAAGAGCAAAAYAAAGAGAAGTAAGTAAACAAATATGTAA 28026

Qy 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28025 CTAAAAAATAAGAGAG 28009

RESULT 15
US-10-995-561-13277/c
; Sequence 13277, Application US/10995561
; Publication No. US20050272054A1
```

```

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13277
; LENGTH: 173995
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-13277

Query Match      13.5%; Score 40.6; DB 10; Length 173995;
Best Local Similarity 55.5%; Pred. No. 18;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAAATATCTTCATTTGAAATAAACAGCTTAACCTCC 77
DB 28061 TACATAGTGAAGTCAAGAGAGAACTTTAAATGTTTACAGATCCAAATATTAAAAATTAA 28002

QY 78 CAAATTCTACTACTAGGGAGTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
DB 28001 CAGAAGTTATTAAAGTAAAGAGCAAAAYAAAGAGAAATGAGTAAAAACAAATATGTAA 27942

QY 138 ATAGAAGTAAACAGAG 154
DB 27941 CTAAAAATAAGAGAG 27925

RESULT 16
US-10-301-480-21248
; Sequence 21248, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21248
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-21248

Query Match      13.2%; Score 39.6; DB 12; Length 483;
Best Local Similarity 51.8%; Pred. No. 2.8;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCAAATGTACTTAACCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 8 CTACTCCATTTGGATCTTAACCTTTGAGATACCAAGAATAATATTGAGCATATCCTTTATAGG 67

QY 134 TCAAATAAGAGTAACCAAGAGTGTTCGAACATGCTGTTAAATAACCCGACAACTTCAAT 193
DB 68 TTAGATACGGAGTTACATCTGACATATAATGTAGTGACATAAGAAAGTAGACTGGAAT 127

QY 194 CACTATAGCTGTAGTAGTGACATCTTCGCAAGGATCCAGAGTAACCA 241
DB 128 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTCAAGGGTAGCAA 175

RESULT 17
US-10-301-480-634657
; Sequence 634657, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 634657
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-634657

Query Match      13.2%; Score 39.6; DB 12; Length 483;
Best Local Similarity 51.8%; Pred. No. 2.8;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCAAATGTACTTAACCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 8 CTACTCCATTTGGATCTTAACCTTTGAGATACCAAGAATAATATTGAGCATATCCTTTATAGG 67

QY 134 TCAAATAAGAGTAACCAAGAGTGTTCGAACATGCTGTTAAATAACCCGACAACTTCAAT 193
DB 68 TTAGATACGGAGTTACATCTGACATATAATGTAGTGACATAAGAAAGTAGACTGGAAT 127

QY 194 CACTATAGCTGTAGTAGTGACATCTTCGCAAGGATCCAGAGTAACCA 241
DB 128 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTCAAGGGTAGCAA 175

RESULT 18
US-10-027-632-93552
; Sequence 93552, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93552
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-93552
```


Query Match 13.2%; Score 39.6; DB 6; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAT 126

QY 194 CACTATAGCTGTAGTAGTGCAATCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTTCAAGGGTAGCA 174

RESULT 19

US-10-027-632-305099
; Sequence 305099, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305099
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305099

Query Match 13.2%; Score 39.6; DB 6; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAT 126

QY 194 CACTATAGCTGTAGTAGTGCAATCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTTCAAGGGTAGCA 174

RESULT 20

US-10-027-632-93552
; Sequence 93552, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93552
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-93552

Query Match 13.2%; Score 39.6; DB 7; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAT 126

QY 194 CACTATAGCTGTAGTAGTGCAATCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTTCAAGGGTAGCA 174

RESULT 21

US-10-027-632-305099
; Sequence 305099, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305099
; LENGTH: 502

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305099

Query Match      13.2%; Score 39.6; DB 7; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCACCAATTTGACTTAACCTAGGAGGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 CTACTCCATTTGGATCTTAATCTTGATACCAAGAAATAATATTGAGCATATCTCTTATAGG 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 134 TCARATAGAGTACCAAGAGTGTTCACATGCTTTAATAAATACCCGACAACTTCAAT 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 TTAGATACGGAGTTACATCTGACATATATATGAGTGTCATAATAGAAAGTAGACTGNAAT 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 194 CACTATAGCTGTAGTAGAGTGATCTCTGCAAGGATCCGAGAGTAACCA 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 127 ACATACAGCATTGAGTAGTAGAAGTTTTCAGGGGTTCAAGGGTAGCAA 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 22
US-09-925-065A-349720
; Sequence 349720, Application US/09925065A
; Publication No. US200401810481
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349720
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349720
```

```
Query Match      13.1%; Score 39.4; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACATGAGAGCTTGAATATTTCTTCAATTTGAATAAAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 TACATAGTGAAGTGAAGAGAACTTTAATGTTTTACAGATCCCAATATTAATAATTTAA 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CAAATTTGACTTAACCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CAGAGTTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGAACAAATATGTAA 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 ATAAGAAGTACACAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 CTAAAAAATAAGAGAG 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 23
US-09-925-065A-349722
; Sequence 349722, Application US/09925065A
; Publication No. US200401810481
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349722
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349722
```

```
Query Match      13.1%; Score 39.4; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACATGAGAGCTTGAATATTTCTTCAATTTGAATAAAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 TACATAGTGAAGTGAAGAGAACTTTAATGTTTTACAGATCCCAATATTAATAATTTAA 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CAAATTTGACTTAACCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CAGAGTTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGAACAAATATGTAA 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 ATAAGAAGTACACAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 CTAAAAAATAAGAGAG 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 24
US-09-925-065A-349720
; Sequence 349720, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349720
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349720
```

```
Query Match      13.1%; Score 39.4; DB 5; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
```


;; TITLE OF INVENTION: in the Human Genome
;; FILE REFERENCE: 108827.137
;; CURRENT APPLICATION NUMBER: US/10/301,480
;; CURRENT FILING DATE: 2002-11-21
;; PRIOR APPLICATION NUMBER: US 10/215,598
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 60/311,695
;; PRIOR FILING DATE: 2001-08-10
;; NUMBER OF SEQ ID NOS: 1226818
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1035189
;; LENGTH: 599
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-301-480-1035189

Query Match 13.1%; Score 39.4; DB 12; Length 599;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 18 TGCTAAGTGAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 77
DB 109 TACATAGTGAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 168
QY 78 CAAATTGTACCTAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 137
DB 169 CAGAAGTTATTAAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 228
QY 138 ATAAGAGTGAAGAGAGAG 154
DB 229 CTAAAGAGTGAAGAGAG 245

RESULT 29
US-10-301-480-1035191
;; Sequence 1035191, Application US/10301480
;; Publication No. US20060057564A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;; FILE REFERENCE: 108827.137
;; CURRENT APPLICATION NUMBER: US/10/301,480
;; CURRENT FILING DATE: 2002-11-21
;; PRIOR APPLICATION NUMBER: US 10/215,598
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 60/311,695
;; PRIOR FILING DATE: 2001-08-10
;; NUMBER OF SEQ ID NOS: 1226818
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1035191
;; LENGTH: 599
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-301-480-1035191

Query Match 13.1%; Score 39.4; DB 12; Length 599;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 18 TGCTAAGTGAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 77
DB 109 TACATAGTGAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 168
QY 78 CAAATTGTACCTAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 137
DB 169 CAGAAGTTATTAAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 228
QY 138 ATAAGAGTGAAGAGAG 154
DB 229 CTAAAGAGTGAAGAGAG 245

RESULT 30
US-09-925-065A-349721
;; Sequence 349721, Application US/09925065A
;; Publication No. US20040181048A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.135
;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; CURRENT FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 349721
;; LENGTH: 602
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-349721

Query Match 13.1%; Score 39.4; DB 4; Length 602;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 18 TGCTAAGTGAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 77
DB 58 TACATAGTGAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 117
QY 78 CAAATTGTACCTAAGTGAAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 137
DB 118 CAGAAGTTATTAAAGTGAAGAGAGAGCTTGAATATTTCTTCAATTTGAATAAACAGCTAACTCC 177
QY 138 ATAAGAGTGAAGAGAG 154
DB 178 CTAAAGAGTGAAGAGAG 194

RESULT 31
US-09-925-065A-349721
;; Sequence 349721, Application US/09925065A
;; Publication No. US20050228172A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.135
;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; CURRENT FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 349721
;; LENGTH: 602
;; TYPE: DNA
;; ORGANISM: Homo sapiens

US-09-925-065A-349721

Query Match 13.1%; Score 39.4; DB 5; Length 602;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATAATCTTCATTGTAATAAAGCAGCTAACTCC 77

Db 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTAAAAATTAA 117

QY 78 CAAATTCCTACTACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137

Db 118 CAGAGTTATTAAAGTAAAGAGCAAAACAAGAGAAATGAGTAAACAAATATGTAA 177

QY 138 ATAAGAAGTAAACACAG 154

Db 178 CTAATAAATAAAGAGAG 194

RESULT 32

US-10-301-480-421781

; Sequence 421781, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE OF INVENTION: in the Human Genome

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/311,695

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 421781

; LENGTH: 608

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-421781

Query Match 13.1%; Score 39.4; DB 12; Length 608;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATAATCTTCATTGTAATAAAGCAGCTAACTCC 77

Db 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTAAAAATTAA 117

QY 78 CAAATTCCTACTACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137

Db 118 CAGAGTTATTAAAGTAAAGAGCAAAACAAGAGAAATGAGTAAACAAATATGTAA 177

QY 138 ATAAGAAGTAAACACAG 154

Db 178 CTAATAAATAAAGAGAG 194

RESULT 33

US-10-301-480-1035190

; Sequence 1035190, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE OF INVENTION: in the Human Genome

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1035190

; LENGTH: 608

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-1035190

Query Match 13.1%; Score 39.4; DB 12; Length 608;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATAATCTTCATTGTAATAAAGCAGCTAACTCC 77

Db 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTAAAAATTAA 117

QY 78 CAAATTCCTACTACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137

Db 118 CAGAGTTATTAAAGTAAAGAGCAAAACAAGAGAAATGAGTAAACAAATATGTAA 177

QY 138 ATAAGAAGTAAACACAG 154

Db 178 CTAATAAATAAAGAGAG 194

RESULT 34

US-10-341-434-82/c

; Sequence 82, Application US/10341434

; Publication No. US20030215835A1

; GENERAL INFORMATION:

; APPLICANT: Origene Technologies

; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes

; FILE REFERENCE: 9U 204 205 R1

; CURRENT APPLICATION NUMBER: US/10/341,434

; CURRENT FILING DATE: 2003-07-18

; PRIOR APPLICATION NUMBER: US 60/348,164

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: US 60/348,119

; PRIOR FILING DATE: 2002-01-15

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 82

; LENGTH: 60461

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (48893)..(49972)

; OTHER INFORMATION:

US-10-341-434-82

Query Match 13.1%; Score 39.4; DB 7; Length 60461;
Best Local Similarity 55.5%; Pred. No. 25;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATAATCTTCATTGTAATAAAGCAGCTAACTCC 77

Db 27844 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTAAAAATTAA 27785

QY 78 CAAATTCCTACTACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137

Db 27784 CAGAGTTATTAAAGTAAAGAGCAAAACAAGAGAAATGAGTAAACAAATATGTAA 27725

QY 138 ATAAGAAGTAAACACAG 154

Db 27724 CTAATAAATAAAGAGAG 27708

RESULT 35

US-10-087-192-331

; Sequence 331, Application US/10087192

; Publication No. US20020182586A1

```
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 331
; LENGTH: 295096
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(295096)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-331.

Query Match      12.9%; Score 38.8; DB 6; Length 295096;
Best Local Similarity 48.9%; Pred. No. 71;
Matches 132; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

1  GGCATTGACAGGCTAAATGCTAAGTGACTACGATGACATGGAATATTTCTTCATTGA 60
175844  GCCATTGCTGAAGACTCTGCTGAGGCTAGATGAAGAGATTCATTAATTCGTTGA 175903
61  AATAAACAGCTAACTCCAAATTTGACCTAACTAGGGGAGTTTATCATGAAGAATGTTT 120
175904  AAAAGAAAGCCTCAAAACAGCTGCTGATAAATTTTGTGTTTATTCAGTTCACTCTTA 175963
121  AATGCTTAATTTTCAATTAAGATGACAGAGTGTGCAACATGCTGT-TAAATACC 179
175964  TGAAGGAATGTTTTTAATAAAGAGGATGAGCTGAGAAAGGAAAAATACAAAATATATGGCT 176023
180  CGACAACTTCAATCACTATAGCTGTAGTAGAGTGCTTCTGCAAGGATCCAGAGTAAC 239
176024  CAAGTATTAAGGGGCATCAAGATGAATGGAGCTGAATCCTATGTTCCAGGAGTTAA 176083
240  CAGTATTTTGGAAATGCAATGTTGAACCGA 269
176084  CAAATAAGGAGTGTTGTAATCTTGAGGCAA 176113

RESULT 36
US-10-357-930-21832
; Sequence 21832, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22160
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_4_5_6_2220_2221_2222_2223_2224
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-22160
```

```
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 331
; LENGTH: 295096
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(295096)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-331.

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy  41  TTGAATATTTCTTCATTGGAATAAACAGAGCTAACTCCAAATTTGCTACCTAAGGGAG 100
Db  2035  TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCATATGATATAACATTATATC 2094
Qy  101  TTATCATGAA-GAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
Db  2095  TATATTATTAAGTAAAGTATTTTTCCTGCTGATTAATGAATATAAAGTAACACAGGGTTA 2154
Qy  160  CAACATGCTGTTAATAAACCCGACAAACTTCAATCACTATA 200
Db  2155  CAGTTGTGCAAAACAAAACATACCTAAATTAACAATA 2195

RESULT 37
US-10-357-930-22160
; Sequence 22160, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22160
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_4_5_6_2220_2221_2222_2223_2224
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-22160
```



```
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24502
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24502

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 41 TTGAATATTTCTTCATTTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAAATTTGTATATAACATTATATC 2094

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGAGTAACCGAGTGTG 159
DB 2095 TATATTATTAAGTAAAGTATTTTGGCTGATTAATGAATATATAAAGTAACAACAGGGTGA 2154

QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAACAAACATACCTCTAAATTACAATA 2195

RESULT 41
US-10-357-930-24579
; Sequence 24579, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24579
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24590

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 41 TTGAATATTTCTTCATTTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAAATTTGTATATAACATTATATC 2094

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGAGTAACCGAGTGTG 159
DB 2095 TATATTATTAAGTAAAGTATTTTGGCTGATTAATGAATATATAAAGTAACAACAGGGTGA 2154

QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAACAAACATACCTCTAAATTACAATA 2195
```

```
RESULT 41
US-10-357-930-24579
; Sequence 24579, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24579
```

```
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24579

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 41 TTGAATATTTCTTCATTTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAAATTTGTATATAACATTATATC 2094

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGAGTAACCGAGTGTG 159
DB 2095 TATATTATTAAGTAAAGTATTTTGGCTGATTAATGAATATATAAAGTAACAACAGGGTGA 2154

QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAACAAACATACCTCTAAATTACAATA 2195

RESULT 42
US-10-357-930-24590
; Sequence 24590, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24590
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24590

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 41 TTGAATATTTCTTCATTTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
DB 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAAATTTGTATATAACATTATATC 2094
```

Qy	101	TTTTATCATGAA-GAATGTTTAAATGCTTAAATTTTCAAAATGAAGATAAACGAGAGTGTG	159
Db	2095	TATATATTAACTGTAAGTATTTTGGCTGANTATGAATAATAAAGTAAACAACAGGGTTA	2154
Qy	160	CAACATCTGTTTAAATAAACCGCAACAACTTCAATCACTATA	200
Db	2155	CAGTTGTGCAAAACCAAAACAAACATACCTTAAATTTACAATA	2195

```

RESULT 43
US-10-357-930-25046
; Sequence 25046, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25046
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-25046

```

Query Match	12.9%;	Score 38.6;	DB 9;	Length 2224;
Best Local Similarity	56.5%;	Pred. No. 10;		
Matches 91;	Conservative 0;	Mismatches 69;	Indels 1;	Gaps 1;
Qy	41	TTGAAATATTCTTCATTTGAAATAAACAAGCTAAACTCCCAAATTTGTACCTAACTAGGGAG	100	
Db	2035	TTGAATTATTTTAAAGCTTAAAAATTCCTGCTGTTTGCAAATTTGTATAAACAATTATATC	2094	
Qy	101	TTTATCATGAR-GAAATGTTTAAATGCTATATTTTCAATTAAGAAGTAACACAGAGTGTG	159	
Db	2095	TATATATTAACTGTAAAGTATTTTTCCTCGAATATGAATATAAAGTAACAACAGGGTTA	2154	
Qy	160	CAACATGCTGTGTTAAATAACCCGACAAACTTCAATCACTATA	200	
Db	2155	CAGTTGTGCAAAACAAAACAAAACATACCTAAATTTACAATA	2195	

RESULT 44

US-10-357-930-27679

; Sequence 27679, Application US/10357930

; Publication No. US20040259086A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel. Robert

```

; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28021,
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28021

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
Db 2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATATATATATATC 2094

Qy 101 TTTATCATGAA-GAAATGCTTTAAATGCTTAATTTTCAAAATAGAGTAACAGAGTGTG 159
Db 2095 TATATTTAATTAAGTAACTGTAAGTATTTTGGCTGATTAATGAAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 46
US-10-357-930-28027
; Sequence 28027, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlögel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28027
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28041

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
Db 2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATATATATATATC 2094

Qy 101 TTTATCATGAA-GAAATGCTTTAAATGCTTAATTTTCAAAATAGAGTAACAGAGTGTG 159
Db 2095 TATATTTAATTAAGTAACTGTAAGTATTTTGGCTGATTAATGAAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 47
US-10-357-930-29041
; Sequence 29041, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlögel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29041
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29041

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
Db 2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATATATATATATC 2094

Qy 101 TTTATCATGAA-GAAATGCTTTAAATGCTTAATTTTCAAAATAGAGTAACAGAGTGTG 159
Db 2095 TATATTTAATTAAGTAACTGTAAGTATTTTGGCTGATTAATGAAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
```

```

; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28021,
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28021

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
Db 2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATATATATATATC 2094

Qy 101 TTTATCATGAA-GAAATGCTTTAAATGCTTAATTTTCAAAATAGAGTAACAGAGTGTG 159
Db 2095 TATATTTAATTAAGTAACTGTAAGTATTTTGGCTGATTAATGAAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 46
US-10-357-930-28027
; Sequence 28027, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlögel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28027
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28041

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
Db 2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATATATATATATC 2094

Qy 101 TTTATCATGAA-GAAATGCTTTAAATGCTTAATTTTCAAAATAGAGTAACAGAGTGTG 159
Db 2095 TATATTTAATTAAGTAACTGTAAGTATTTTGGCTGATTAATGAAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 47
US-10-357-930-29041
; Sequence 29041, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlögel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29041
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29041
```

```
Db      2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 48
US-10-027-632-86622/c
; Sequence 86622, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86622
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(552)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86622

Query Match      12.7%; Score 38.2; DB 6; Length 552;
Best Local Similarity 56.0%; Pred. No. 7.2;
Matches 70; Conservative 1; Mismatches 54; Indels 0; Gaps 0;

Qy      108 TGAAGAAATGTTAAATGCTTAATTTTCAATAAGAGTAACCGAGTGTTCGAACATGC 167
Db      282 TAAATAATGTTAAAGGGTAATTTACAAAAAATGAATAATAAAGTGGACATGAAGT 223
Qy      168 TGTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCTCAAGGA 227
Db      222 GATTAATAACAGAAATCCAGAGCTATCATTATTGAYGTATGATATCTTCAATTTTAAAAA 163
Qy      228 TCCCA 232
Db      162 GGACA 158

RESULT 49
US-10-027-632-86623/c
; Sequence 86623, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86623
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(552)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86623

Query Match      12.7%; Score 38.2; DB 6; Length 552;
Best Local Similarity 56.0%; Pred. No. 7.2;
Matches 70; Conservative 1; Mismatches 54; Indels 0; Gaps 0;

Qy      108 TGAAGAAATGTTAAATGCTTAATTTTCAATAAGAGTAACCGAGTGTTCGAACATGC 167
Db      282 TAAATAATGTTAAAGGGTAATTTACAAAAAATGAATAATAAAGTGGACATGAAGT 223
Qy      168 TGTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCTCAAGGA 227
Db      222 GATTAATAACAGAAATCCAGAGCTATCATTATTGAYGTATGATATCTTCAATTTTAAAAA 163
Qy      228 TCCCA 232
Db      162 GGACA 158

RESULT 50
US-10-027-632-316419/c
; Sequence 316419, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316419
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(552)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-316419
```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 22:44:54 ; Search time 118 Seconds
(without alignments)
3544.714 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1 ggcattgacaggtaaatgc.....atctgctaggtaggaagtat 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 886355 seqs, 697127050 residues

Total number of hits satisfying chosen parameters: 1772710

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New.*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	38.6	12.9	863	8	US-11-266-748A-249055	Sequence 249055,
2	37.8	12.6	1000	8	US-11-266-748A-201627	Sequence 201627,
3	37.8	12.6	1000	8	US-11-266-748A-281554	Sequence 281554,
4	37.8	12.6	1000	8	US-11-266-748A-308194	Sequence 308194,
5	37.8	12.6	1000	8	US-11-266-748A-390539	Sequence 390539,
6	37.8	12.6	1000	8	US-11-266-748A-481257	Sequence 481257,
7	37	12.3	6857	8	US-11-266-748A-23772	Sequence 23772, A
8	36.8	12.3	692	8	US-11-266-748A-16908	Sequence 16908, A
9	36.4	12.1	415	8	US-11-266-748A-360446	Sequence 360446,
10	36.4	12.1	415	8	US-11-266-748A-443825	Sequence 443825,
11	35.6	11.9	578	8	US-11-266-748A-362773	Sequence 362773,
12	35.6	11.9	578	8	US-11-266-748A-446152	Sequence 446152,
13	35.6	11.9	1000	8	US-11-266-748A-403961	Sequence 403961,
14	35.6	11.9	1000	8	US-11-266-748A-475007	Sequence 475007,
15	35.6	11.9	2205	8	US-11-266-748A-31158	Sequence 31158, A
16	35.6	11.9	130263	8	US-11-266-748A-58551	Sequence 58551, A
17	35.2	11.7	352	8	US-11-266-748A-257564	Sequence 257564,
18	35.2	11.7	352	8	US-11-266-748A-278222	Sequence 278222,
19	35.2	11.7	352	8	US-11-266-748A-318081	Sequence 318081,
20	35.2	11.7	1421559	8	US-11-266-748A-28208	Sequence 28208, A
21	34.6	11.5	633	6	US-10-471-571A-3251	Sequence 3251, Ap
22	33.6	11.2	1108	8	US-11-266-748A-353636	Sequence 353636,
23	33.6	11.2	1108	8	US-11-266-748A-437015	Sequence 437015,
24	33.6	11.2	1908	8	US-11-266-748A-188843	Sequence 188843,

33.6	11.2	10389	8	US-11-266-748A-30896	Sequence 30896, A
33.4	11.1	1000	8	US-11-266-748A-200842	Sequence 200842,
33.4	11.1	1000	8	US-11-266-748A-201296	Sequence 201296,
33.4	11.1	1278	8	US-11-217-529-81496	Sequence 81496, A
33.2	11.1	1146	8	US-11-266-748A-77181	Sequence 77181, A
33.2	11.1	1146	8	US-11-266-748A-129992	Sequence 129992,
32.8	10.9	594	8	US-11-217-529-5685	Sequence 5685, Ap
32.8	10.9	109669	7	US-11-021-837-41	Sequence 41, Appl
32.8	10.9	111597	8	US-11-266-748A-22665	Sequence 22665, A
32.8	10.9	247877	8	US-11-266-748A-22663	Sequence 22663, A
32.8	10.9	1071650	8	US-11-266-748A-22664	Sequence 22664, A
32.6	10.9	4857	6	US-10-517-441-424	Sequence 424, App
32.4	10.8	1000	8	US-11-266-748A-224134	Sequence 224134,
32.4	10.8	1000	8	US-11-266-748A-292677	Sequence 292677,
32.4	10.8	1000	8	US-11-266-748A-344106	Sequence 344106,
32.4	10.8	1000	8	US-11-266-748A-404470	Sequence 404470,
32.4	10.8	1000	8	US-11-266-748A-475516	Sequence 475516,
32.4	10.8	1039	8	US-11-266-748A-71950	Sequence 71950, A
32.4	10.8	1039	8	US-11-266-748A-124761	Sequence 124761,
32.4	10.8	8391	8	US-11-266-748A-24100	Sequence 24100, A
32.4	10.8	8391	8	US-11-217-529-166179	Sequence 166179,
32.2	10.7	1218	8	US-11-266-748A-82454	Sequence 82454, A
32.2	10.7	1218	8	US-11-266-748A-111321	Sequence 111321,
32.2	10.7	4388	8	US-11-266-748A-135265	Sequence 135265,
32.2	10.7	4388	8	US-11-266-748A-30890	Sequence 30890, A
32.2	10.7	26683	7	US-11-021-837-26	Sequence 26, Appl
32	10.7	1113	8	US-11-217-529-82444	Sequence 82444, A
32	10.7	1924	8	US-11-266-748A-73868	Sequence 73868, A
32	10.7	1924	8	US-11-266-748A-126679	Sequence 126679,
32	10.7	7126	8	US-11-266-748A-58818	Sequence 58818, A
32	10.7	7126	8	US-11-266-748A-28338	Sequence 28338, A
32	10.7	140167	8	US-11-266-748A-60770	Sequence 60770, A
31.8	10.6	527	8	US-11-266-748A-307049	Sequence 307049,
31.8	10.6	637	8	US-11-216-545-3893	Sequence 3893, Ap
31.8	10.6	693	8	US-11-266-748A-35109	Sequence 35109, A
31.8	10.6	932	8	US-11-266-748A-19166	Sequence 19166, A
31.6	10.6	4722	8	US-11-266-748A-28340	Sequence 28340, A
31.6	10.5	458	8	US-11-266-748A-92792	Sequence 92792, A
31.6	10.5	458	8	US-11-266-748A-145603	Sequence 145603,
31.6	10.5	1000	8	US-11-266-748A-281058	Sequence 281058,
31.6	10.5	1000	8	US-11-266-748A-307698	Sequence 307698,
31.6	10.5	1000	8	US-11-266-748A-398345	Sequence 398345,
31.6	10.5	1000	8	US-11-266-748A-469391	Sequence 469391,
31.6	10.5	1219	6	US-10-509-131-27	Sequence 27, Appl
31.6	10.5	1409	7	US-11-218-305-23785	Sequence 23785, A
31.6	10.5	1807	8	US-11-216-545-6502	Sequence 6502, Ap
31.6	10.5	5275	8	US-11-266-748A-61192	Sequence 61192, A
31.4	10.5	654	8	US-11-266-748A-412734	Sequence 412734,
31.4	10.5	663	8	US-11-266-748A-49205	Sequence 49205, A
31.4	10.5	663	8	US-11-266-748A-209620	Sequence 209620,
31.4	10.5	663	8	US-11-266-748A-234110	Sequence 234110,
31.4	10.5	892	8	US-11-266-748A-182511	Sequence 182511,
31.4	10.5	892	8	US-11-266-748A-243056	Sequence 243056,
31.4	10.5	5189	8	US-11-266-748A-27190	Sequence 27190, A
31.4	10.5	5189	8	US-11-266-748A-30394	Sequence 30394, A
31.4	10.5	5967	8	US-11-266-748A-32555	Sequence 32555, A
31.4	10.5	8622	8	US-11-319-952-21	Sequence 21, Appl
31.2	10.4	901	8	US-11-266-748A-354433	Sequence 354433,
31.2	10.4	901	8	US-11-266-748A-437812	Sequence 437812,
31.2	10.4	913	8	US-11-266-748A-8033	Sequence 8033, Ap
31.2	10.4	1194	8	US-11-217-529-81713	Sequence 81713, A
31.2	10.4	1563	8	US-11-217-529-75623	Sequence 75623, A
31.2	10.4	2503	8	US-11-266-748A-61671	Sequence 61671, A
31.2	10.4	55266	8	US-11-266-748A-24926	Sequence 24926, A
31	10.3	846	8	US-11-266-748A-296367	Sequence 296367,
31	10.3	1000	8	US-11-266-748A-198313	Sequence 198313,
31	10.3	1427	8	US-11-266-748A-193857	Sequence 193857,
31	10.3	1585	8	US-11-266-748A-181087	Sequence 181087,
31	10.3	5405	8	US-11-266-748A-29238	Sequence 29238, A
31	10.3	150000	8	US-11-266-748A-23591	Sequence 23591, A
31	10.3	684973	8	US-11-266-748A-32013	Sequence 32013, A
30.8	10.3	9001	6	US-10-517-441-790	Sequence 790, App


```

; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-481257

Query Match      12.6%; Score 37.8; DB 8; Length 1000;
Best Local Similarity 57.0%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 52;

QY 8 ACAGGCTAAATGCTAGTGAAGTACGATGAAGACTTGAAGATATTTCTTCATTGAAATAAAC 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ACTGCCCTACAACAAGGCGAGTTCTCTGAACCACTGATTATTTTCAATTATTCATCA 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 68 AGCTAACTCCCAATTTGCTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCT 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 ATTATTAATTTTCAAAATTTATACCTTAAAAATGTTAGTTGGTAAATTAAGACATAGTACTAGTT 389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 T 128
Db 390 T 390

RESULT 7
US-11-266-748A-23772/c
; Sequence 23772, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23772
; LENGTH: 6857
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-23772

Query Match      12.3%; Score 37; DB 8; Length 6857;
Best Local Similarity 55.9%; Pred. No. 1.7; Indels 1; Gaps 1;
Matches 90; Conservative 0; Mismatches 70;

QY 41 TTGAATATTTCTTCATTGTTGAATTAACAGCTAACTCCCAATTTGCTACCTAATAGGGAG 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCATAATTTCTTTTAACTATATC 2175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAAACAGAGTGTG 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2174 TATATTTAACTGTAAGTATTTTTCCTGCTGATTATGAATATAAAGTAAACACAGGGTTA 2115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 160 CAACATGCTGTAAATAACCCGACAAACTTCAATCACTATA 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-390539

Query Match      12.6%; Score 37.8; DB 8; Length 1000;
Best Local Similarity 57.0%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 52;

QY 8 ACAGGCTAAATGCTAGTGAAGTACGATGAAGACTTGAAGATATTTCTTCATTGAAATAAAC 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 ACTGCCCTACAACAAGGCGAGTTCTCTGAACCACTGATTATTTTCAATTATTCATCA 672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 68 AGCTAACTCCCAATTTGCTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCT 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 ATTATTAATTTTCAAAATTTATACCTTAAAAATGTTAGTTGGTAAATTAAGACATAGTACTAGTT 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 T 128
Db 611 T 611

RESULT 6
US-11-266-748A-481257
; Sequence 481257, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 481257
; LENGTH: 1000
```

```

, PRIOR APPLICATION NUMBER: EP 04105479.2
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: EP 04105482.6
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: EP 04105483.4
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: EP 04105507.0
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: EP 04105485.9
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: EP 04105484.2
, PRIOR FILING DATE: 2004-11-03
, PRIOR APPLICATION NUMBER: US 60/662,276
, PRIOR FILING DATE: 2005-03-14
, PRIOR APPLICATION NUMBER: US 60/700,293
, PRIOR FILING DATE: 2005-07-18
, NUMBER OF SEQ ID NOS: 483996
, SOFTWARE: PatentIn version 3.3
, SEQ ID NO 360446
, LENGTH: 415
, TYPE: DNA
, ORGANISM: Homo Sapiens
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (12)..(12)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (66)..(66)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (73)..(73)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (131)..(131)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (376)..(376)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (394)..(394)
, OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-360446

Query Match 12.1%; Score 36
Best Local Similarity 51.2%; Pred. No
Matches 85; Conservative 0; Mismatch

QY 46 ATATTTCTTCATTGTAATAATAACAGCTAAAC
||| ||| ||| ||| |||
Db 182 ATATAAATCAGTCACACAAATGTATAAT

QY 106 CATGAAGAAATGTTTAAATGCTTAAATTT
||| ||| ||| ||| |||
Db 242 AAATCAGCTGTTTAAATCTATATATT

QY 166 GCTGTTTAAATAACCCGACAAACTTCAAT
||| ||| ||| ||| |||
Db 302 TTTTATAAGGAGTGTATAAAGATATGCC

RESULT 10
US-11-266-748A-443825/c
; Sequence 443825, Application US/112667488/c
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl

```

```

; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 443825
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40)..(40)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (285)..(285)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (343)..(343)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (350)..(350)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (404)..(404)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-443825

```

```

Query Match 12.1%; Score 36.4; DB 8; Length 415;
Best Local Similarity 51.2%; Pred. No. 1.3;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 46 ATATCTTCTTTTGAATAAACAAGTCTCCCAAAATTTACCTAAGTGGGAGTTTAT 105
Db 234 ATATAATCAGTCAACAATGTAATGGACTAAAGTTACCAAGTTAGATTGAATTAAT 175

Qy 106 CATGAAGAATGTTTAAATGCTTAATTTCAATAAAGAAGTAACACAGAGTGTTCACAT 165
Db 174 AAAATCCAGCTGTTTAAATCTATATTTAAATATGAAGGTACAAAAGGTTGAAAGTT 115

Qy 166 GCTGTTAAATTAACCGCAAACTCAATCACTATAGCTGTAGTAGA 211
Db 114 TTTTAAGGAGAGTGTAAAGATATGCCAGTATACATACTAGA 69

```

```

RESULT 11
US-11-266-748A-362773
; Sequence 362773, Application US/11266748A

```

```

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 362773
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-362773

Query Match 11.9%; Score 35.6; DB 8; Length 578;
Best Local Similarity 50.6%; Pred. No. 2.2;
Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 36 AAGACTTGAATATTTCTTCAATTTGAATAAACAAGTAACTCCCAAAATTTACCTAAGT 95
Db 409 AGGACATGGAGTCTACTGTATTTAAATATAGAGCTCATCTCAAAATTTCTATATGTAAG 468

Qy 96 GCGAGTTTATCATGAAGAATGTTTAAATGCTTAATTTCAATAAAGAAGTAACCAAGT 155
Db 469 AGCATTTTCTTTGAAAACATGTATATGGATAATAATAATTAATTAATTAATTAAT 528

Qy 156 GTTGCAACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGT 205
Db 529 GAAATATCAGACATAGAAATCATAGCATCTATTAAACAACTTTGCTTT 578

```

```

RESULT 12
US-11-266-748A-446152/c
; Sequence 446152, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9

```



```

; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31158
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-31158

Query Match      11.9%; Score 35.6; DB 8; Length 2205;
Best Local Similarity 50.6%; Pred. No. 3.1;
Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 36 AAGCTTGAATATCTTCTTCAATTAACAGCTAATCCCAAAATGTACCTAAGT 95
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2008 AGGACATGGAGTCTACTGTATTTAAATATAGAGCTCATCTCAAAATCTATATGTTAAAG 2067
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 96 GGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAATAAGAGTAAACAGAGT 155
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2068 AGCATTTCTTTGAAACATGATATATGGATAAATAATTAATCTACTTCTTAAATGTTT 2127
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 156 GTTCCAAATGCTCTTTAAATTAACCCGACAACTTCAATCACTATAGCTGT 205
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2128 GAAATATCAGACATAGAAATCATAGCATCTATTTAAACAACTTTCGCTTT 2177
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 16
US-11-266-748A-58551/c
; Sequence 58551, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31158
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Homo Sapiens

```

```

; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58551
; LENGTH: 130263
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-58551

Query Match      11.9%; Score 35.6; DB 8; Length 130263;
Best Local Similarity 49.5%; Pred. No. 8;
Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATTTCTTCAATTTGAAATATAACAGCTAATCCCAAT 83
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAATAAGA 143
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 144 AGTAACGAGAGTCTTGCACATCTGTTAAATTAACCCGACAACTTCAATCACTATAGCT 203
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 35994 AGTCCACAGAGAAATATCCAAAGTGAAGAGCTAAGAAATAGCAGTATATAAAAT 35995
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 204 GTAGTA 209
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 35934 GAAATA 35929
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 17
US-11-266-748A-257564/c
; Sequence 257564, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 257564
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:

```



```
NAME/KEY: misc_feature
LOCATION: (67)..(67)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117)..(117)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-257564

Query Match      11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 GAACAAATATCAGTAATCTCTTTGTTCTAAACAAAAATTCATAATTTATTATACATTT 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 TTAATGCTTAATTTTCAAAATAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAAC 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 TAAATATTATATTGTTTCAAATGTTGTAGTGGGCAATAAATCATAAAGNGATACAAC 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGATCCAGAGT 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 CAGTCGAAATTAGACTTTTAAAGATTTACTCTCAANGACATTGCAACAACTCCAAAGT 95
```

```
RESULT 18
US-11-266-748A-278222/c
; Sequence 278222, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 278222
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: n is a, c, g, or t
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (67)..(67)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117)..(117)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-278222
```

```
Query Match      11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 GAACAAATATCAGTAATCTCTTTGTTCTAAACAAAAATTCATAATTTATTATACATTT 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 TTAATGCTTAATTTTCAAAATAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAAC 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 TAAATATTATATTGTTTCAAATGTTGTAGTGGGCAATAAATCATAAAGNGATACAAC 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGATCCAGAGT 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 CAGTCGAAATTAGACTTTTAAAGATTTACTCTCAANGACATTGCAACAACTCCAAAGT 95
```

```
RESULT 19
US-11-266-748A-318081
; Sequence 318081, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 318081
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (192)..(192)
```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (236)..(236)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (279)..(279)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (286)..(286)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (303)..(303)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-318081
```

```
Query Match 11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 59 GAATAAAGAGTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 GAACAAATATCAGTAATCTCTTTGTTCTTAACAAAAATTCATAATTTATACATTT 140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 119 TTAATGCTTTAAATTTCAAAATAAGAGTAACCGAGAGTGTTCACACATGCTGTTAAATPAC 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 TAAATATTATTATGTTTCAAAATCTGTGTAGTGGGCAATAAATCATAAAGNGATACAAC 200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 179 CCGACAAACTTCATCTACTATAGCTGTAGTAGATGCTTTCGAAGGATCCGAGGT 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 CAGTCGAATTAGACTTTTAAAGATTTACTCTCAANGACATTCGAACAACTCCAAAGT 258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 20

```
US-11-266-748A-28208
; Sequence 28208, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28208
; LENGTH: 1421559
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (414394)..(414394)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (836909)..(836909)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1022504)..(1022504)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1041105)..(1041105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1122667)..(1122667)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-28208
```

```
Query Match 11.7%; Score 35.2; DB 8; Length 1421559;
Best Local Similarity 53.7%; Pred. No. 18;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 15 AAATGCTAGTGAAGTACGATGAAGACTTGAATATTTCTTCAATTTGAATAAAGAGCTAAC 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289724 AAAAAAGAAAAACAAAAGAGATCAATGGAATCAATTTGAACTTGATATAAATAGAAATA 289783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 75 TCCCAATTTACTACTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTT 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289784 GCCAAATAGTAGGTGCATAGGAGAAATATCACAAGGTGCATGAGTAATTTTCCAAATTT 289843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 135 CAAATAAGAGTAACC 150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289844 TGAGGCTTAGTAACC 289859
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 21

```
US-10-471-571A-3251/c
; Sequence 3251, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3251
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3251

Query Match 11.5%; Score 34.6; DB 6; Length 633;
Best Local Similarity 56.6%; Pred. No. 4.2;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 105 TCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTACCAAGAGTGTTCACACA 164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 389 TTATTATTATAAATTTCCAACTTAGCTTAGCTTACGATTCAGAAATAAAAAAGTATTAAACA 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 165 TGCTGTTAAATAACCCGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCAT 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 CCAGAAATATTGCACACATATTACAAATCACCATGTTTCTCTAATTGAAT 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 22

```
US-11-266-748A-353636/c
; Sequence 353636, Application US/11266748A
```

```
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US 60/700,293
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 353636
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1060)..(1060)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-353636

Query Match 11.2%; Score 33.6; DB 8; Length 1108;
Best Local Similarity 46.6%; Pred. No. 9;
Matches 108; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATAATTTCTTCAATTTGAAATAAACAAGCTAACTCCCAAAATGTACC 88
DB 257 TGCATATGAGTTTCGATTAACTGATAATTTGGAATTTGGGTCCTAAGATATAAAC 198
QY 89 TAACTAGGGAGTTTATCATGAGAGAAATGTTTAAATGCTTAAATTTTCAATAAGAGTAA 148
DB 197 AGAATGGAGAAATTAATGGAGAGTAACCTTTTCATAGCTGTATTATATAAAGGGTGCACAC 138
QY 149 CCAGAGTGTTCACACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGT 208
DB 137 ATTTGACAGCTCAGACACTCTTGTCAAGAGCCTACTAGCAAGTGTCAAGGTGTGGG 78
QY 209 AGAGTGCAATCTGCAAGGATCCAGAGTAACCCAGATTAATTTTGGAAATGCAATG 260
DB 77 CAACGTCTTCTTCGAGGCTCCAGAAAGAACCTTATTCTTGGTGAAGGAAG 26

RESULT 23
US-11-266-748A-437015
; Sequence 437015, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
```


Db 1127 TTCAATGAACTACATAGCAGCTAAATAGAGAGCTTCCAAAATGTATACAGCAATTGGAT 1068
Qy 101 TTTATCATGAAGAAATGTTAAATGCTTAATTTTCAAAATAGAGAGTAACAGAGTGTGTC 160
Db 1067 CTAACCTAAGGAGAGCTTCAGAAATTTTAAATTAAGATAAAGAGTGTTCACCTGTGAAA 1008
Qy 161 AACATGCTGTTAAATACCCGACAACTTCATCACTATAGCTGTA 206
Db 1007 AATTACTATTAAATAATTTTAAATTAAGAAATTTTAAACTATA 962

RESULT 30

US-11-266-748A-129992
; Sequence 129992, Application US/11266748A
; Publication No. US20060134663A1

GENERAL INFORMATION:

; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)

; CURRENT APPLICATION NUMBER: US/11/266,748A

; CURRENT FILING DATE: 2005-11-03

; PRIOR APPLICATION NUMBER: EP 04105479.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105483.4

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105507.0

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105485.9

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105484.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: US 60/662,276

; PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293

; PRIOR FILING DATE: 2005-07-18

; NUMBER OF SEQ ID NOS: 48396

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 129992

; LENGTH: 1146

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-11-266-748A-129992

Query Match 11.1%; Score 33.2; DB 8; Length 1146;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 41 TTGAATATCTTCATTGTAATAAAGCAAGCTTCCCAATTTGCTAACTAGGAG 100

Db 20 TTCAATGAACTACATAGCAGCTTAAATAGAGAGCTTCCCAATTTGCTAACTAGGAG 79

Qy 101 TTTATCATGAAGAAATGTTAAATGCTTAATTTTCAAAATAGAGAGTAACAGAGTGTGTC 160

Db 80 CTAACCTAAGGAGAGCTTCAGAAATTTTAAATTAAGATAAAGAGTGTTCACCTGTGAAA 139

Qy 161 AACATGCTGTTAAATACCCGACAACTTCATCACTATAGCTGTA 206

Db 140 AATTACTATTAAATAATTTTAAATTAAGAAATTTTAAACTATA 185

RESULT 31

US-11-217-529-5685

; Sequence 5685, Application US/11217529

; Publication No. US20060099612A1

GENERAL INFORMATION:

; APPLICANT: SUNTORY LIMITED

; APPLICANT: NAKAO, YOSHIHIRO

; APPLICANT: NAKAMURA, NORIHIISA

; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5685
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5685

Query Match 10.9%; Score 32.8; DB 8; Length 594;
Best Local Similarity 56.5%; Pred. No. 13;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 81 ATTGTACTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140

Db 206 ACTGGGACTATACAGAATTTCAAACCAAGAAATTTTCAGCAGCTAAATTTGACGAAGA 265

Qy 141 AGAAGTAACAGAGTGTGCAACATGCTGTTAAATACCCGACAACT 188

Db 266 TAAATCATCTGAGAGTAACGGCATAAAGAAAAATTTACTCGAAAACT 313

RESULT 32

US-11-021-837-41

; Sequence 41, Application US/11021837

; Publication No. US20060140972A1

GENERAL INFORMATION:

; APPLICANT: Alm, Richard

; APPLICANT: Manning, Paul

; APPLICANT: McLaughlin, Robert

; APPLICANT: McCormack, Kathleen

; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides

; FILE REFERENCE: 101324-US

; CURRENT APPLICATION NUMBER: US/11/021,837

; CURRENT FILING DATE: 2004-12-23

; PRIOR APPLICATION NUMBER: US 60/533534

; PRIOR FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/600680

; PRIOR FILING DATE: 2003-08-31

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 41

; LENGTH: 109669

; TYPE: DNA

; ORGANISM: Staphylococcus saprophyticus

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (46850)..(46850)

; OTHER INFORMATION: n is a, c, g, or t

US-11-021-837-41

Query Match 10.9%; Score 32.8; DB 7; Length 109669;
Best Local Similarity 52.1%; Pred. No. 44;
Matches 73; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 2 GCATTTGACAGGCTAAATGCTTAAGTACTAGCAGTAAGAGCTTTGAAATTTCTTCAATTTGAA 61

Db 30576 GCAAAATCAAATAAATTAACAACCTATTCAAATAATAGTTAGCTGTTTTTAAATTCG 30635

Qy 62 ATAAAGAGCTTAACCTCCCAAAATTTGCTTAAGTACTAGGAGTTTATCATGAAGAAATGTTTA 121

Db 30636 AAAATGGCGGAAATGGTTGGTTGGAATAGTAAGGACAGAGAAAAATAAATATTTC 30695

Qy 122 AATGCTTTAATTTTCAAATAA 141


```
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 292677
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-292677

Query Match      10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAATTTGTAACCTAAGGAGTTTATCATGAAGAAATGT 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 902 GAAACAATATCAGTAATCTCTTTGTTCTAAACAAAATTCATAATTTATTTATACATTT 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 TTAATGCTTAATTTTCAAATAAGAAAGTAACCAAGAGTGTGCAACATGCTGTTAAATAAC 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 842 TAAATATTTATTTGTTTCAAAATGTTGTAGTGGGCAATAAATCATAAAGATACAC 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 CCGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCAGAGT 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 782 CAGTCGAAATAGACTTTTAAAGAATTAATCTCTCAAAGACATTTGCCAACAACTCCAAAGT 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 39
US-11-266-748A-344106
; Sequence 344106; Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 344106
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-404470

Query Match      10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAATTTGTAACCTAAGGAGTTTATCATGAAGAAATGT 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 902 GAAACAATATCAGTAATCTCTTTGTTCTAAACAAAATTCATAATTTATTTATACATTT 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 TTAATGCTTAATTTTCAAATAAGAAAGTAACCAAGAGTGTGCAACATGCTGTTAAATAAC 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 842 TAAATATTTATTTGTTTCAAAATGTTGTAGTGGGCAATAAATCATAAAGATACAC 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 CCGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCAGAGT 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 782 CAGTCGAAATAGACTTTTAAAGAATTAATCTCTCAAAGACATTTGCCAACAACTCCAAAGT 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 39
US-11-266-748A-344106
; Sequence 344106; Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
```

Db 782 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCAAAGT 725

RESULT 41

US-11-266-748A-475516
; Sequence 475516, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 475516
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-475516

Query Match 10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 59 GAAATAACAGCTAACTCCCAAAATTGTACCTTAAGTGGGAGTTTATCATGAAGAAATGT 118
Db 99 GAAACAATATCAGTAATCCTCTTTGTTCTTAACAAAAATTCATAATTTATACATTT 158
QY 119 TTAATGCTTAATTTTCAAATAAGAGTAACAGAGTGTTCGAACATGCTGTTTAAATAAC 178
Db 159 TAAATATTATATTGTTTCAAATCTTGTAGTGGGCAATAAATCATAAAGAGATACAAC 218
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTGCAAGATCCCAAGT 236
Db 219 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCAAAGT 276

RESULT 42

US-11-266-748A-71950/c
; Sequence 71950, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 71950
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-71950

Query Match 10.8%; Score 32.4; DB 8; Length 1039;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 59 GAAATAACAGCTAACTCCCAAAATTGTACCTTAAGTGGGAGTTTATCATGAAGAAATGT 118
Db 957 GAAACAATATCAGTAATCCTCTTTGTTCTTAACAAAAATTCATAATTTATACATTT 898
QY 119 TTAATGCTTAATTTTCAAATAAGAGTAACAGAGTGTTCGAACATGCTGTTTAAATAAC 178
Db 897 TAAATATTATATTGTTTCAAATCTTGTAGTGGGCAATAAATCATAAAGAGATACAAC 838
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTGCAAGATCCCAAGT 236
Db 837 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCCAAGT 780

RESULT 43

US-11-266-748A-124761
; Sequence 124761, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124761
; LENGTH: 1039
; TYPE: DNA

```

; ORGANISM: Homo Sapiens
US-11-266-748A-124761

Query Match      10.8%; Score 32.4; DB 8; Length 1039;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATGTTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
DB 83 GAAACAAATATACGTAATCCTCTTTGTTCTAAACAAAATTCATATTTATTTATACATTT 142
QY 119 TTAATGCTTAAATTTTCAAAATGAAGTAACACAGAGTGTGCAACATCTGTTAAATAAC 178
DB 143 TAAATATATATGTTTCAATGTTGTTAGTGGGCAATTAATCATAAAGAGATACAAAC 202
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCCTGCAAGGATCCCAAGT 236
DB 203 CAGTCGAATTAGACTTTTAAAGRATTACTCTCAAAGACATTCGCAACAACTCCCAAGT 260

RESULT 44
US-11-266-748A-24100/c
; Sequence 24100, Application US/112666748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24100
; LENGTH: 4862
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-24100

Query Match      10.8%; Score 32.4; DB 8; Length 4862;
Best Local Similarity 48.9%; Pred. No. 27;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATGTTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
DB 4781 GAAACAAATATACGTAATCCTCTTTGTTCTAAACAAAATTCATATTTATTTATACATTT 4722
QY 119 TTAATGCTTAAATTTTCAAAATGAAGTAACACAGAGTGTGCAACATCTGTTAAATAAC 178
DB 4721 TAAATATATATGTTTCAATGTTGTTAGTGGGCAATTAATCATAAAGAGATACAAAC 4662
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGTGCAATTCCTGCAAGGATCCCAAGT 236
DB 4661 CAGTCGAATTAGACTTTTAAAGRATTACTCTCAAAGACATTCGCAACAACTCCCAAGT 4604
```

```

RESULT 45
US-11-217-529-166179/c
; Sequence 166179, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166179
; LENGTH: 8391
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166179

Query Match      10.8%; Score 32.4; DB 8; Length 8391;
Best Local Similarity 58.2%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 40 CTTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTCTACTACTAGGGGA 99
DB 5463 CTTTATATATTCTTAAATAAAATATAAAGCAATTAATAATAATAATAATAATAATAA 5404
QY 100 GTTTATCATGAGAAATGTTTAAATGCTTAAATTTTCA 137
DB 5403 CAATATAATAATAATAATAATAATAATAATAATAACAA 5366

RESULT 46
US-11-266-748A-22572
; Sequence 22572, Application US/112666748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22572
; LENGTH: 31857
; TYPE: DNA
```

```
; ORGANISM: Homo Sapiens
US-11-266-748A-22572

Query Match      10.8%; Score 32.4; DB 8; Length 31857;
Best Local Similarity 55.3%; Pred. No. 42;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 52 TTCATTTGAAATTAACAGCTAACCTCCAAATTTGTAACCTAACTAGGGGAGTTTATCATGAA 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26990 TGCATTTTACCTACATACAAACCTACAAATGTGTAACCTAACTAGGGGAGTTTATCATGAA 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 112 GAAATGTTTAAATGCTTAATTTTCAATTAAGATGCTACCAAGCTGTTGCAACAT 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27050 GAAAAATGTTGTACATGAATGCTCATAGAAGCAGTATTATTAGAATAGTCAAAAAT 27103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 47
US-11-266-748A-82454/c
; Sequence 82454, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82454
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-82454

Query Match      10.7%; Score 32.2; DB 8; Length 1218;
Best Local Similarity 59.1%; Pred. No. 22;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 42 TGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAATTTGACCTAACTAGGGGAGT 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 343 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAATATATATATACCGAGGAGT 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 TTCAAAATTAATTTATATTCACAGGCTTGCTTTT 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 48
US-11-266-748A-111321/c
; Sequence 111321, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82454
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-82454

Query Match      10.7%; Score 32.2; DB 8; Length 1218;
Best Local Similarity 59.1%; Pred. No. 22;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 42 TGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAATTTGACCTAACTAGGGGAGT 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 343 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAATATATATATACCGAGGAGT 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 TTCAAAATTAATTTATATTCACAGGCTTGCTTTT 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 49
US-11-266-748A-135265
; Sequence 135265, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82454
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-82454
```

```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135265
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-135265

Query Match      10.7%; Score 32.2; DB 8; Length 1218;
Best Local Similarity 59.1%; Pred. No. 22;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 42 TGAATAATTCCTCATTTGCAATAAACAAGCTAACTCCCAAAATGTACCTAACTAGGGGAGT 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAAAATATATATACCAGGAGCAGT 935
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 936 TTCAAATTAATATATTACAGGGCTTGCTTTT 968
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 50
US-11-266-748A-30890/c
; Sequence 30890, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30890
; LENGTH: 4388
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-30890

Query Match      10.7%; Score 32.2; DB 8; Length 4388;
Best Local Similarity 59.1%; Pred. No. 30;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 42 TGAATAATTCCTCATTTGCAATAAACAAGCTAACTCCCAAAATGTACCTAACTAGGGGAGT 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3210 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAAAATATATATACCAGGAGCAGT 3151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3150 TTCAAATTAATATATTACAGGGCTTGCTTTT 3118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: July 18, 2006, 00:52:40
Job time : 128 secs

```

This Page Blank (uspto)